Journal of Experimental Psychology

ARTHUR W. MELTON, Editor
Department of Psychology, University of Michigan
Ann Arbor, Michigan

DAVID A. GRANT, Associate Editor
University of Wisconsin

WILLIAM K. ESTES, Associate Editor
Indiana University

Consulting Editors

NORMAN H. ANDERSON, *University of California, Los Angeles*
E. JAMES ARCHER, *University of Wisconsin*
FRED ATTNEAVE, III, *University of Oregon*
CLETUS J. BURKE, *Indiana University*
PAUL M. FITTS, *Ohio State University*
FREDERICK C. FRICK, *Massachusetts Institute of Technology*
FRANK A. GELDARD, *University of Virginia*
JAMES J. GIBSON, *Cornell University*
CLARENCE H. GRAHAM, *Columbia University*
HAROLD W. HAKE, *University of Illinois*
LLOYD G. HUMPHREYS, *University of Illinois*
ARTHUR L. IRION, *Tulane University*
HOWARD H. KENDLER, *New York University*
DONALD B. LINDSLEY, *University of California, Los Angeles*
KENNETH MACCORQUODALE, *University of Minnesota*
QUINN MCNEMAR, *Stanford University*
L. STARLING REID, *University of Virginia*
KENNETH W. SPENCE, *State University of Iowa*

CONTENTS

Learning a Motor Task Under Varied Display Conditions: N. B. GORDON 65
The Decremental Effects of Partial Reinforcement During Acquisition of the Conditioned Eyelid Response: L. E. ROSS 74
The Difference Threshold of the Magnitude of Visual Velocity: B. B. BRANDALISE AND R. M. GOTTSDANKER 83
Information and Autokinetic Movement: W. EDWARDS 89
Finger Span: Ratio Scale, Category Scale, and JND Scale: S. S. STEVENS AND G. STONE .. 91
Paired-Associates Learning as a Function of Percentage of Occurrence of Response Members (Reinforcement): A. E. GOSS, C. H. MORGAN, AND S. J. GOLIN 96
Prediction of Sequential Two-Choice Decisions From Event Runs: D. C. NICKS 105
Performance as a Joint Function of Amount of Reinforcement and Interreinforcement Interval: G. COLLIER AND M. SISKEL, JR. 115
Spontaneous Regression and Recovery in a Sequence of Discrimination Periods: D. MCCONNELL 121
The Role of Reinforcement and Nonreinforcement in an "Apparent Frustration Effect": A. R. WAGNER 130
Erratum 136

American Psychological Association
Vol. 57, No. 2 February 1959

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc., at Prince and Lemon Sts., Lancaster, Pa. The subscription rate per volume is $8.00, or $16.00 annually. Single copies are $1.50. Subscriptions, orders, address changes, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This Journal publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects, except when the latter are specifically oriented toward the extension of behavior theory. Experimental psychometric studies and studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology.

Normally, articles exceeding 20 printed pages cannot be accepted. Within this limit, the piecemeal experiment-by-experiment reporting of psychological research is discouraged, and the reporting of the data and conclusions of substantial segments of research efforts is encouraged. Specifically, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) *must* be presented in a single article. Special provision is made for the publication of Supplementary Reports and Replication Reports in articles of not more than 1.5 printed pages (see Editor's Note, this Journal, 1957, 53, 1–2).

Address all articles submitted for publication to the Editor, Arthur W. Melton, Department of Psychology, University of Michigan, Ann Arbor, Michigan. **Manuscripts *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc., as described in the *Publication Manual of the American Psychological Association*.** When in doubt about practices of the Journal, authors should examine a recent issue.

Articles are published in the order of their receipt, except in rare circumstances. Authors are supplied with 50 free offprints without covers. All of the cost of an author's alterations in galley proof is charged to the author. Priority in publication is given to articles whose authors assume the full cost of publication. In 1959 the publication cost is $30 per printed page. Authors of priority publications receive no free offprints.

To save printing costs, supplementary material to articles in this Journal are deposited with the American Documentation Institute. Copies may be obtained by ordering the Document Number given in the footnote reference to ADI. Orders should be sent to ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. Orders *must* be accompanied by checks or money orders payable to the Chief, Photoduplication Service, Library of Congress.

ARTHUR C. HOFFMAN, Managing Editor; HELEN ORR, Promotion Manager; SADIE J. DOYLE, Editorial Assistant

Second-class postage paid at Lancaster, Pa.

Send address changes to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must reach the Subscription Office by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

© 1959 by the American Psychological Association, Inc.

Journal of
Experimental Psychology

VOL. 57, NO. 2 FEBRUARY, 1959

LEARNING A MOTOR TASK UNDER VARIED DISPLAY CONDITIONS [1,2]

NORMAN B. GORDON [3]

New York University

Continuous perceptual-motor (tracking) tasks have come to present considerable conceptual difficulty to the theorist interested in extending accepted learning principles to them. There are at least two basic causes for this difficulty. First, it is extremely difficult to break a continuous task down into conventionally identifiable stimulus and response components. Second, even where it is possible to do so, it is found that tasks may be similar in terms of specific stimuli and response but differ in terms of the temporal and sequential relationships among these elements. It is often found that apparently similar tasks produce significantly different learning when compared in transfer designs.

Thus, Lincoln (1954) found that a group of *Ss* originally trained with verbal knowledge of results in a rate-of-hand-crank-turning task learned the rate of movement significantly better than groups given visual as well as verbal knowledge of results in training. It has been observed that there is little intertask transfer between pursuit and compensatory tracking (Andreas, Green, & Spragg, 1954). In a comparison between instrument and contact flight performance, Ritchie and Michael (1955) found that while instrument flight was a more difficult task to learn, it produced higher transfer to contact flight than prior training on contact produced for instrument flight. In mirror-tracing experiments, mirror reversal after initial training produced negative transfer if *Ss* ignored what Cook (1947) called "kinesthetic memory" of the path and attended to visual cues. Bilodeau (1952) found strong evidence that two similar motor tasks

[1] This article is based upon a dissertation submitted to the Department of Psychology in the Graduate School of Arts and Science at New York University in partial fulfillment of the requirements for the Ph.D. degree. The author wishes to express his appreciation to James N. Farr, Richard S. Christie, and Douglas H. Fryer, who successively sponsored the thesis.

[2] The research was a part of Project 20-OS-14 in the Human Engineering Department of the United States Naval Training Device Center, Port Washington, New York, and the author is deeply grateful for the assistance and support given by the Center to this research.

[3] Now at Albert Einstein College of Medicine, Yeshiva University, Bronx 61, New York.

produced different kinds of learning as a result of differences in the character of the feedback information. In each of the above instances, the same responses were required in the various versions of the tasks used. There is a strong suggestion that the cause of differential transfer is probably the fact that the differences in external stimuli result in different learning by the *S*s. This possibility offers a better explanation than the conventional attempt to ascribe differential transfer to "task difficulty."

In recent work by Poulton an attempt has been made to deal with the problems of the kinds of adaptations made in learning continuous complex tasks. Poulton (1952a, 1952b) has pointed out that tasks by their very nature range from those in which *S* is continuously dependent on external stimuli ("open tasks") to those where *S* learns the entire sequence of required responses and achieves independence of signals from the environment ("closed tasks"). In a more recent experiment, Poulton (1957) has compared various degrees of intermittency in external stimulus information with continuous information during a pursuit tracking task. He points out that where stimulus information is obliterated for brief periods of time, *S*'s error rises as if he were not tracking at all, but that practice with the eyes closed in the use of kinesthetic cues could improve the matching of a simple input when the display was invisible. Thus, data from a variety of sources suggest that complex tracking tasks may vary considerably in the kinds of adaptation required by *S* to perform them.

The purpose of this study is to test the hypothesis that differential learning and retention may result when tasks requiring identical responses differ both quantitatively and qualitatively in the stimulus dimension. Specifically, it is hypothesized that stimulus changes affect the kind and amount of learning and retention needed to perform the task. For example, dialing a particular number on the telephone when the numbers are visible is a task which differs greatly from one where *S* must dial the same number under blindfolded conditions, yet the required motor tasks are identical. An individual who has learned the latter task will probably be able to perform the former without additional learning. This is probably not true if the former is learned first. The real difference is the manner in which *S* "acquires the stimulus."

Some tasks provide meager external perceptual aspects and hence require a rote-like memorization of the sequences of responses or of stimuli. Tasks which are perceptually richer require less learning because the quality and number of stimuli available during performance facilitate getting to the correct responses. To study the consequences of such variables, four tasks were designed requiring identical responses but differing in the following way. The first situation, called *pattern memory* (PM), requires *S* to track an unseen target, the path of which has been presented in advance. The second task, called *pursuit tracking* (PT), requires tracking of a seen target with a follower. The third task, called *response memory* (RM), requires tracking of a seen target with an unseen follower. The fourth task, called *compensatory tracking* (CT), requires canceling out follower movement with respect to a fixed target. By experimental design all possible transfer combinations among the four tasks were observed.

METHOD

Experimental Design

After pretraining, 16 groups of *Ss* were given nine training trials, four groups to each task (PM, PT, RM, and CT). After training, one of the four groups for each original task was transferred to the same task or one of the other three tasks. The second period also consisted of nine trials for all groups. The design is a simple factorial transfer experiment. Each individual trial was 45 sec. long and was followed by 1 min. of rest. Each set of three trials was separated by 3-min. rest periods; and the two periods of nine trials were separated by 5-min. rest periods.

Apparatus

The experimental apparatus consisted of portions of an experimental tracking device developed by the U. S. Naval Training Device Center[4] for the purpose of demonstrating the effect of different control-display relationships on performance. Two different displays and a single set of controls were used (see Fig. 1).

Controls.—Two rotary hand-crank controls were used in this experiment. One hand crank controlled a vertical axis and the other a horizontal axis on the display panel. Each control moved in the same way. When centered it was in the zero position. The control was spring loaded so that when released it returned to the zero position. It was possible to move the controls approximately 160° to the right or left of center. The hand cranks were connected to precision potentiometers to provide signals which operated the displays and provided scoring information.

Displays.—Two displays of information were used. The *combined display* consisted of a $9\frac{1}{2} \times 9\frac{1}{2}$-in. vertically placed, slightly curved white screen constructed of polished Lucite and rendered translucent by application of flat white paint to the surface away from *S*. Two circles of light representing target and follower were projected on the screen by projectors whose movement was controlled by the program unit. The larger circle ($\frac{3}{4}$-in. diameter) was the target and the smaller ($\frac{3}{8}$-in. diameter) was the follower.

[4] The experimental apparatus used in this research is that developed as Project 9-U-72, "Human Engineering Principles Demonstrator" in the U. S. Naval Training Device Center Laboratory.

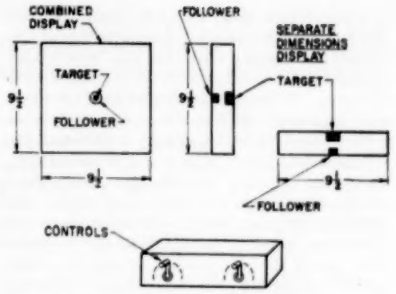

FIG. 1. Experimental tracking apparatus.

Either of the stimuli could be turned off without affecting the actual system movement. The combined display was used for the pattern memory, pursuit, and compensatory tracking tasks. The *separate dimension display* consisted of two manometer strips, one vertical and the other horizontal. Each was $9\frac{1}{2}$-in. long. A target indicator moved back and forth on one side of each strip and directly opposite the follower indicator. The follower was painted red (see Fig. 1). The separate dimension display was used for response memory to eliminate the possibility of seeing the tracing of the two-dimensional path of the target and to force *S* to concentrate on a more "rote-like" learning of responses. The center of the display was 46 in. from the floor corresponding to average eye level of seated *Ss*. The controls were placed 24 in. away from the display on a a table directly in front of *S*.

Program mechanism.—This unit combines inputs from the controls and from an electromechanical cam on which task information is stored. The output of the program mechanism is transmitted mechanically to the displays and resulted in both target information and *S*-produced display changes.

Scoring.—Scoring was accomplished by continuously comparing target position with *S*'s responses. The output of the scoring mechanism consisted of a varying voltage which represented the magnitude of the distance between target and follower, and was recorded on Brush recording paper. For scoring purposes the Brush record was read once each second. The error score consisted of the sum of 45 readings for each trial.

General Procedure

All *Ss* were familiarized with the device by pretraining in the operation of the controls

in conjunction with the combined display. The pretraining consisted of explanation of the control and display operation and then actual operation of the follower to trace a few simple patterns. In the case of *Ss* who used the separate dimension display either during first or second experimental periods, additional instruction was given prior to experimentation. They were informed that the new display gave information equivalent to that from the combined display, but in two separate parts, and were asked to manipulate the controls and observe the relationship of control movements with the follower movements. Specific pretraining test patterns were given. No attempt was made to explain to *S* that the required responses were the same for the tasks they would practice (most *Ss* practiced two different tasks).

Experimental Instructions and Conditions

Pattern memory.—This task was performed with the combined display. The *S* was told that the target would trace out a sort of clover-leaf pattern. An overlay of the pattern was placed over the display and *S* was allowed to observe the target light following the tracing. The *S* was then told that he would have to remember the path taken by the target, as well as the rate at which it moved, as it would not be visible during his performance. His task was defined as that of "tracking an unseen target." He was further told that he would be given knowledge of results at the end of each trial, and allowed to see a tracing of the total path again before subsequent trials. He was allowed to study the pattern again during a 45-sec. observation of the target movement, and the experiment was begun. The target light remained off, but the follower light was visible. At the end of each trial *S* was shown a tracing of his path compared with that of the target. He was verbally informed of rate errors. At the end of Trials 3 and 6, he was again allowed to observe one cycle of operation of the target prior to proceeding.

Response memory.—This task used the separate dimension display. After pretraining on the display (described above) *S* was told that the target, which consisted of vertical and horizontal pointers, would move

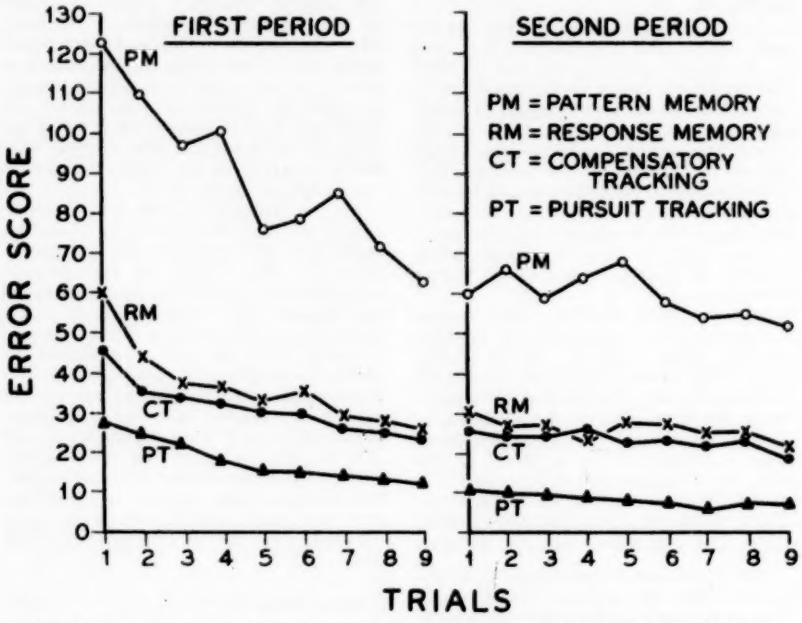

FIG. 2. Error scores on the four experimental tasks of four groups of 40 *Ss* each during training and of one control group of 10 *Ss* each during the second or transfer period.

TABLE 1

ANALYSES OF VARIANCE OF LOG MEAN ERROR SCORES

Source of Variance	*df*	First Period		Second Period					
		Trials 7–9		Trials 1–3		Trials 4–6		Trials 7–9	
		MS	*F*	*MS*	*F*	*MS*	*F*	*MS*	*F*
First-task effects (A)	3	3.631	57.635**	.206	1.421	.217	1.064	.324	2.746
Second-task effects (B)	3	.050	.794	4.695	32.379**	4.132	20.255**	5.137	43.534**
A × B	9	.063	1.703	.145	3.816*	.204	4.163*	.118	2.950*
Within groups	144	.037		.038		.049		.040	
Total	159								
Bartlett's χ^2 test [a]	15	23.98		21.29		24.11		23.71	

Note.—The interaction error term was used as the denominator in determining the *F* ratios because the design is a Type II or components of variance design in which the levels of each factor were randomly selected.

[a] Required for significance at .05 level: 24.996.

* $P \leqq .01$, against within-group error.

** $P \leqq .01$, against interaction error.

or rest independently of each other. The *Ss*' task was to operate the controls so as to match the target position from moment to moment. The follower pointers, however, remained at rest, so that *S* had to remember the appropriate position of the controls for various positions of the target pointers. Knowledge of results was given verbally after each trial by *E*. At the end of the Trials 3 and 6, *S* was again allowed to manipulate the follower pointers prior to proceeding.

Pursuit tracking.—The *Ss* were given the following instructions for this task which used the combined display:

"In this task the target you see in front of you will start to move on the display. You will be required to manipulate your controls in such a way that you keep the follower superimposed on the target during the trial. It will not be too fast or too slow. Try to work as smoothly as possible. Do not be abrupt in your movements. You will be able to tell how well you are doing by observing how closely the target and follower overlap during performance."

The target described a clover-leaf type of path identical to that of the pattern memory task.

Compensatory tracking.—The combined display was used. Instructions for this task were identical to those of pursuit tracking. The task differed in that the target remained stationary during the task, and the follower was moved by both *S* and the apparatus. The *S*'s task consisted of nulling follower movement and of attempting to keep the follower aligned with the target.

Subjects

The 160 male *Ss* were selected at random from three classes at the U. S. Merchant Marine Academy, Kings Point, New York. The *Ss* were randomly assigned to one of 16 experimental conditions and were scheduled for a testing hour.

RESULTS

Figure 2 summarizes the results of the experiment insofar as the comparison of performances on the four tasks is concerned. The left-hand side of Fig. 2 is a plot of the error scores of four groups of 40 *Ss* during nine trials (first set). The right-hand side of the figure shows the error scores of four groups of 10 *Ss* continuing the same task practiced in the first set, during the second set of nine trials. All groups reached the approximate asymptote of performance after nine trials. The groups show statistically reliable performance differences with the exception of Re-

sponse Memory and Compensatory Tracking.

The error scores were subjected to analysis of variance at four points. Each point represents the average error score during three trials. The points selected were Trials 7–9 during the first period and Trials 1–3, 4–6, and 7–9 of the second period. The error scores were converted to logarithms to satisfy the requirement of homogeneity of variance.

Table 1 is a summary of the analyses. During training, the groups practicing first tasks differed significantly. At the same time, those groups which will in the second period practice different tasks do not vary significantly.

During the transfer (second) period, groups do not differ significantly with respect to the over-all effect of first-task training on second-task performance (the first-task main effect during the second period). However, the second-task main effects become highly significant. Significant interactions are found at the three points tested during the second period, giving evidence of differential transfer effects.

Figure 3 is a presentation of the effect on tasks in the second or transfer period of type of task practiced during the first period. One group practicing each set continued the same task in the second period; all of the others practiced new tasks. From these comparisons, it is quite evident that learning resulting from practice

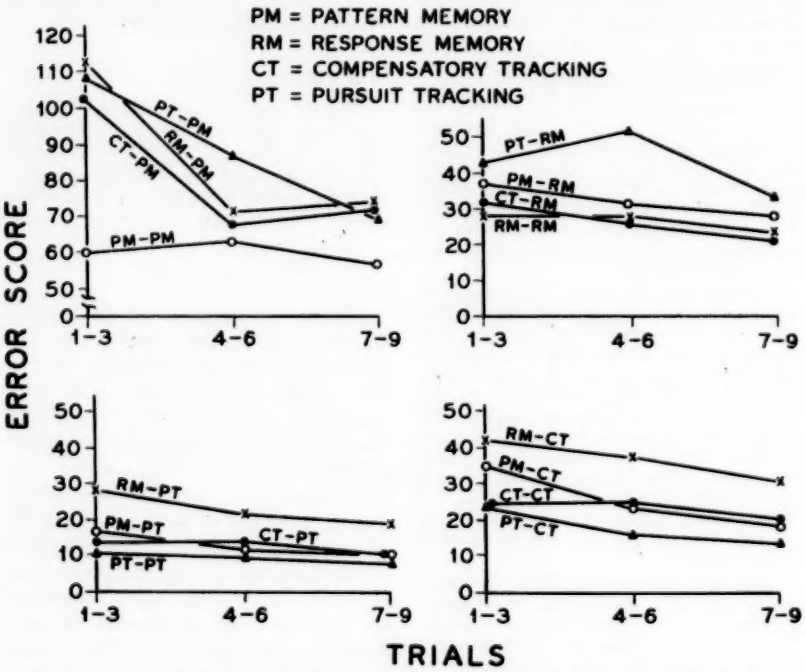

FIG. 3. Effects during the second (transfer) period of type of task practiced during the first period on second tasks. Each group consists of 10 *Ss*.

TABLE 2

PERCENTAGE TRANSFER SCORES RESULTING FROM DIFFERENTIAL FIRST-PERIOD TRAINING

Second-Period Task	First-Period Task											
	Trials 1–3 (Second Task)				Trials 4–6 (Second Task)				Trials 7–9 (Second Task)			
	PM	PT	RM	CT	PM	PT	RM	CT	PM	PT	RM	CT
PM	(100)	−14*	−21*	0*	(100)	42	81	95	(100)	68	64	64
PT	65*	(100)	−17*	80	81	(100)	31*	71	86	(100)	50*	88
RM	59*	29*	(100)	93	80	−18*	(100)	103	83*	60*	(100)	103
CT	35*	102	−16*	(100)	107	157*	30*	(100)	114	133	42*	(100)

* Differences between control and transfer groups significant at .05 level.

in the first period is differentially effective for new tasks depending upon which task follows. This is further confirmed in Table 2 which is a summary of the analyses of differential transfer effects and is based on an application of Formula 3 in Gagné, Foster, and Crowley (1948). From Table 2 and Fig. 3 it can be seen that amount of transfer of training is not uniformly the same from a given first task to all second tasks. Furthermore, if Fig. 2 is examined along with Table 2, it can be seen that amount of transfer does not appear to be directly related to level of performance achieved during training.

DISCUSSION

The four tasks of this experiment constituted variations of the display conditions under which a motor task was to be performed. Performance of each of the conditions resulted in varying degrees of proficiency and varying transfer to each of the other conditions. The following is an attempt to comprehend some of the phenomena that were observed.

Improvement in such perceptual-motor tasks takes place when there is an observable improvement in timing and speed of response. In the present experiment, *S* is directly dependent upon external stimulus information provided by displays. If the information is not sufficient, *S* becomes dependent upon direct (response produced) or indirect (*E* provided) knowledge of results. Implied in the concept of knowledge of results is some permanent adaptation (or a form of learning) applicable to the task. This can consist of learning the perceptual pattern of external stimuli in such a way that required responses can be predicted ahead for a period of time (Poulton's "perceptual anticipation") or the kinesthetically mediated rote-like learning of a set of required responses, or some combination of the two.

In the present experiment, the displays were so arranged as to emphasize the requirement for different kinds of adaptation by *S*. Pursuit tracking provided optimum information, which included both knowledge of results and ample knowledge of what to do at the right time. Pattern memory provided a high degree of information concerning external stimuli but knowledge of results was available only indirectly. Compensatory tracking provided a high degree of knowledge of results, but little information that would aid in predicting external stimuli. In response memory the precise pattern of external stimuli was broken up through use of the separated dimension display and knowledge of results was only indirectly available. These tasks thus varied in the

the degree to which *S* was receiving information necessary to produce a sequence of appropriate responses. Figure 2 is a graphic illustration of the effect of these information differences. Even though there is a general tendency for all groups to improve, approximately the same degree of separation between groups is maintained throughout. The pattern-memory groups suffer on two accounts, both in terms of delayed knowledge of the pattern as well as knowledge of the proper timing of responses (since target information is not available). The other tasks all provide timing information, but obviously differ in terms of the usefulness of this information to the task at hand. It is interesting to note that performance on compensatory tracking and response memory, though not differing significantly, produce greatly differing transfer results.

The major purpose of the experiment was to determine if the differences in information availability between the tasks would, in fact, result in different learning patterns by *S* as a consequence of the kind of adaptation required by a particular task. It was hypothesized that learning as inferred on the basis of transfer results would be a function of both the amount and the kind of learning required by the training task. In other words, the amount of transfer to a new task was expected to be dependent on the learning acquired during the training task even though this learning was not demonstrated in terms of proficiency during performance of the training task. Thus, pursuit tracking was expected to demonstrate relatively less transfer than other tasks. It does in fact transfer poorly to pattern memory and response memory. Its efficiency as a training task for compensatory tracking (over 100% transfer) gives evidence that the task of target-follower alignment was facilitated in pursuit tracking by the high degree of anticipatory information presented by target movement in a well-organized perceptual pattern. The *S* was able to effect a reduction in delay very readily. In transfer to compensatory tracking this manipulative ability, plus familiarity with the approximate sequence and magnitude of the required responses as a temporal sequence, undoubtedly was responsible for the high degree of facilitation. The same approximate familiarity with response sequences was not sufficiently precise in transfer to the "memory" tasks where *S* is required to learn much more about the nature of the relationships between available stimuli and required responses.

Compensatory tracking and pattern memory ranked highest in transfer value and hence gave evidence of producing the greatest amount of transferable learning. Each of these tasks provides less stimulus information than pursuit tracking. Each in its own way requires *S* to make an additional adaptation to that required by pursuit tracking. The pattern memory task requires *S* to remember the entire pattern or tracing of the target path and, in addition, to remember the correct timing of responses in order to perform adequately. This is the most difficult task in terms of achieved proficiency. However, the resulting adaptations lead to high positive transfer to the other tasks. Here again degree of proficiency on the training task was no indication of potential transfer value. Compensatory tracking leaves out almost entirely the anticipatory information provided in the pursuit tracking task. The *S* in this task receives no knowledge of the over-all pattern of the target but only discrete changes as the follower is displaced from the center of the display. The only way that *S* could learn this task with zero error was either to remember the entire pattern of responses by kinesthetic means or to develop such a high degree of proficiency in response to the discrete changes in the display that he could respond by immediately producing the required movement. Here again transfer is uniform and high and apparently bears no relationship to the degree of proficiency during training. Response memory is the poorest transfer producer. Even though *S* in response memory was extremely dependent upon

experimenter-provided knowledge of results, the intermittency and extreme rote-like character of the information he acquired made what he learned but marginally useful in the other, more well-organized tasks.

In view of the differential effectiveness of training on one task for transfer to the other tasks, it is concluded that *S* learns different things as a function of the particular relationships between stimuli and responses in a given task. While one can only speculate concerning the kinds of learning that take place in perceptual-motor tasks (and the data in the present experiment can only be taken as indicative of a range of possibilities), it is clear that complex perceptual-motor tasks require highly specialized and differential adaptations. Such variations in adaptation can even occur in tasks requiring identical responses and where the stimuli in an objective sense are highly similar. Learning must be some function of what *S* must acquire and retain in order to meet task requirements. This learning may or may not be appropriate to new similar tasks.

Summary

This study is concerned with an examination of the hypothesis that learning in perceptual-motor (tracking) tasks is a function of the nature of the information contained in the stimulus inputs to *S*. An experiment was performed in which identical two-handed tracking responses were required in the context of four variations of stimulus inputs. The variations consisted of pursuit, compensatory, pattern memory, and response memory tasks. A 4×4 factorial design was used to study learning and all possible transfer combinations between sets of two tasks. Nine training and nine transfer trials were used. Results give evidence supporting the hypothesis that what is learned, and hence transferred, is a function of amount and kind of stimulus information available in training. Amount of transfer of training was a function of both the transferable learning that took place during training and the relevance of the learning for a given transfer task.

References

Andreas, B. G., Green, R. F., & Spragg, S. D. S. Transfer effects between performance on a following tracking task and a compensatory tracking task. *J. Psychol.*, **1954, 37,** 173–183.

Bilodeau, E. A. Transfer of training between tasks differing in degree of physical restriction of imprecise responses. *USAF Hum. Resour. Res. Cent. Res. Bull.*, No. 52–40, 1952.

Cook, T. W. Interference and mirror position. *Canad. J. Psychol.*, 1947, **3,** 116–126.

Gagné, R. J., Foster, H., & Crowley, M. E. The measurement of transfer of training. *Psychol. Bull.*, 1948, **45,** 97–130.

Lincoln, R. S. Learning a rate of movement. *J. exp. Psychol.*, 1954, **47,** 465–470.

Poulton, E. C. Perceptual anticipation in tracking with two-pointer and one-pointer displays. *Brit. J. Psychol.*, 1952, **43,** 222–229. (a)

Poulton, E. C. The basis of perceptual anticipation in tracking. *Brit. J. Psychol.*, 1952, **43,** 295–302. (b)

Poulton, E. C. On the stimulus and response in pursuit tracking. *J. exp. Psychol.*, 1957, **53,** 189–194.

Ritchie, M., & Michael, A. L. Transfer between instrument and contact flight training. *J. appl. Psychol.*, 1955, **39,** 145–149.

(Received March 28, 1958)

Journal of Experimental Psychology
Vol. 57, No. 2, 1959

THE DECREMENTAL EFFECTS OF PARTIAL REINFORCEMENT DURING ACQUISITION OF THE CONDITIONED EYELID RESPONSE [1]

LEONARD E. ROSS

State University of Iowa

Studies involving classical defense conditioning of the human eyelid response have compared the acquisition performance resulting from 100% (continuous) and 50% (partial, intermittent) reinforcement schedules. Two studies, one by Humphreys (1939) and one by Grant, Riopelle, and Hake (1950), failed to find continuous reinforcement superior to 50% reinforcement. Although Humphreys did not find a significant difference between the two conditions, evidence has been presented (Grant & Hake, 1951) that indicates this may have been due to the inclusion in the data of sensitized "Beta" responses occurring in the first 200 msec. following the onset of the CS. Grant, Riopelle, and Hake's negative findings involved the use of response magnitude as the behavioral measure, the magnitude of CR's being greater with 50% reinforcement than with 100% reinforcement.

Later studies (Grant & Hake, 1951; Grant & Schipper, 1952; Grant, Schipper, & Ross, 1952; Gynther, 1957; Reynolds, 1958) which have avoided the inclusion of "Beta" responses and used frequency as the response measure, have consistently found higher levels of conditioning with 100% reinforcement. The evidence thus indicates that, under these conditions, 50% partial reinforcement produces a decrement in conditioning performance when compared to that obtained by continuous reinforcement.

Study of this decremental effect of partial reinforcement has been handicapped by the small numbers of acquisition trials that have been given. The primary concern in the majority of cases has been in the effect of reinforcement schedule upon resistance to extinction and as a result the acquisition periods have been limited in length. Only Gynther's (1957) study, which was not concerned with extinction, has involved as many as 100 acquisition trials. If learning (habit growth) occurs only on reinforced trials in aversive conditioning there will be, in the case of 50% partial reinforcement, only half as many increments of habit strength compared with 100% reinforcement for the same number of trials. To investigate the decremental effect of nonreinforced trials without the confounding of differential numbers of reinforcements, a 50% reinforcement group should have had double the number of trials received by the 100% reinforcement group at the point of comparison. This can be achieved by making comparisons in terms of reinforced trials rather than by trials alone.

Since a maximum of 50 reinforced acquisition trials has been given, study of the effects of 50% reinforcement for a period at and beyond the asymptote of learning is not possible with the available data. If some

[1] This study is based on a thesis submitted to the Graduate College of the State University of Iowa in partial fulfillment of the requirements for the M.A. degree. The writer is indebted to Kenneth W. Spence for advice and assistance throughout the course of the investigation.

inhibitory factor continues to develop on nonreinforced trials the performance curve should drop after this point. It is possible, on the other hand, that during prolonged training the performance of a 50% group might approach that of a 100% group.

Switching from partial reinforcement to continuous reinforcement and vice versa offers another means of investigating the effects of nonreinforced trials. Switches to partial reinforcement should provide information regarding the rate of development of factors depressing performance and, similarly, changes from partial to continuous reinforcement should result in evidence relating to the dissipation of inhibitory factors. Further data of value would be gained from determination of the final levels of performance following switching.

One study which involved a switch in reinforcement conditions was that of Kimble, Mann, and Dufort (1955). In the course of investigating "classical" and "instrumental" eyelid conditioning (UCS on every trial and UCS omitted when a response was made to the CS) shifts were made from one reinforcement condition to the other. This in effect resulted in mean switches from 100% to less than 10% reinforcement and from 40% to 100% reinforcement. The group shifted from continuous reinforcement showed a small but significant drop in performance while the group shifted to continuous reinforcement jumped over 40% within a block of 10 trials. These results are of limited relevance to the present problem due to the small N (six per group) and the fact that the amount of change in reinforcement varied from S to S depending upon the individual's performance.

The present experiment was designed to provide data on the effect of 50% partial reinforcement upon acquisition of the conditioned eyelid response over a longer period than has heretofore been investigated, and also to provide data on the effects of switching from continuous to partial and from partial to continuous schedules of reinforcement.

Method

Subjects.—Seventy-seven women and 35 men from a course in introductory psychology served as *Ss*. Two women and one man who gave CR's to test trial presentations of the CS alone were discarded as were four women and two men who exceeded the criterion used to exclude voluntary responders. This criterion, as in previous studies (e.g., Reynolds, 1958; Spence & Taylor, 1951) was the occurrence of more than 50% CR's with a latency of less than 300 msec. Three additional *Ss* were discarded, one due to *E* error and two due to equipment failure. The 100 remaining *Ss* were assigned to five groups at random except that the proportion of women to men was fixed at a ratio of seven women to three men.

Apparatus and method of recording.—The *Ss* were seated in an adjustable dental chair which was located in a semisoundproof room illuminated by a shielded 7½-watt bulb and separated by a third intervening room from that housing the recording apparatus and stimulus controls. The equipment for recording eyelid closure, which involved a potentiometer-polygraph system, has been described elsewhere (Spence & Taylor, 1951).

The CS was an increase in brightness of a 2¼-in diameter circular milk-glass disk, located 51 in. in front of and slightly above the eye level of the *S*, from less than .004 ml. to .23 ml. The UCS was a 50-msec. 2.0-lb./sq.-in. air puff delivered through a .062-in. diameter orifice to the right eye by a 110-v., 60-cycle AC-operated solenoid valve. The onsets and durations of the CS and the UCS were controlled by means of Hunter-Brown (1949) decade-type electronic interval timers.

Conditioning procedure.—Each *S* was given instructions to blink once to the ready signal and then to look at the disk on the wall, remaining as relaxed as possible, until the light went out. Following the reading of the instructions each *S* received three presentations of the CS alone, to check for any initial conditioned response tendency to the

light, and one presentation of the UCS alone. Following Trial 80 the headset adjustment for each *S* was checked. This involved an interruption of less than 30 sec. No additional lights were turned on in the experimental room, dim illumination from an adjoining hallway providing sufficient light. The interval between the ready signal (the word "Ready") and the CS onset was 2, 3, or 4 sec. randomly varied. Intertrial intervals of 15, 20, and 25 sec., averaging 20 sec. and arranged by fixed schedule, were used.

A CR was recorded whenever the record showed a deflection of 1 mm. or more in the interval 200–500 msec. following the onset of the CS.

Experimental design.—Two reinforcement conditions, 100% and 50% were employed. The 50% reinforcement condition consisted of half reinforced and half nonreinforced trials presented in a prearranged sequence within which neither type of trial occurred more than twice in succession. A reinforced trial involved a 500-msec. CS-UCS interval while on nonreinforced trials the UCS followed the CS onset at 2400 msec., an interval at which McAllister (1953a; 1953b) has shown that very little or no conditioning occurs. In both types of trials the duration of the CS extended 50 msec. beyond the onset of the UCS, both stimuli terminating together.

Five groups of 20 *S*s (14 women and 6 men) were conditioned. Two groups received the same type of trials throughout conditioning, Group C being given continuous reinforcement for 200 trials and Group P receiving a 50% partial reinforcement schedule for 220 trials. Three other groups were switched from one condition of reinforcement to a new reinforcement schedule. Group P 40 C was switched from 50% partial to continuous reinforcement at Trial 40, receiving a total of 120 trials, while two other groups, C 20 P and C 100 P, were switched from continuous to 50% partial reinforcement after Trials 20 and 100, respectively. The total number of trials received by Group C 20 P was 120. Group C 100 P received a total of 200 trials.

Results

Partial versus continuous reinforcement.—Acquisition curves for Groups C and P in terms of the percentage of conditioned responses occurring in successive blocks of 10 trials are shown in Fig. 1. Inspection of the curves reveals that the continuous reinforcement schedule produced a consistently higher level of performance. The Mann-Whitney **U** test was used to test the difference between Group C and Group P during Trials 1–20, 21–40, 41–60, 101–120, and 181–200. The difference was significant in each case with two-tailed

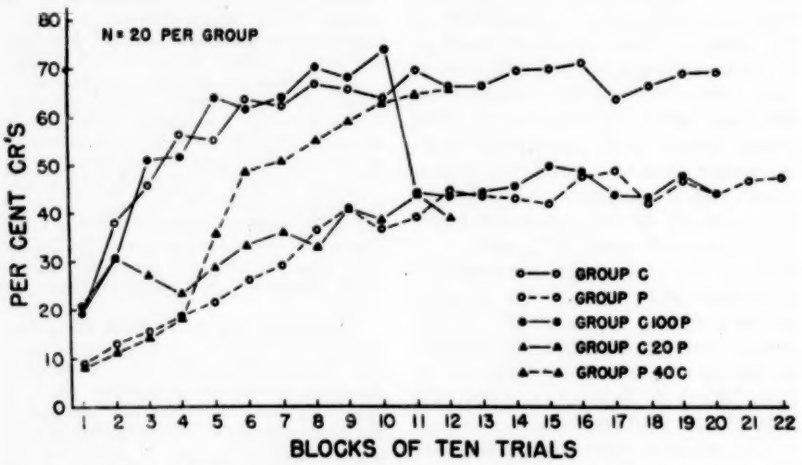

FIG. 1. Mean percentage of CR's in successive blocks of 10 trials.

P values of .019, .004, .006, .016, and .006, respectively. It should be noted that the middle three blocks of trials tested (21–40, 41–60, 101–120) were those immediately following points where switches in reinforcement schedule occurred for the other three groups and that at these points Curves C and P were significantly different.

Unless otherwise stated, Mann-Whitney **U** tests were used in the comparisons which follow. A two-tailed P value is given in each case.

Partial to continuous reinforcement.—Figure 1 also presents the curve for Group P 40 C. No significant difference was found between this group and Group P on Trials 21–40 ($P = .401$). Following the switch in reinforcement schedule the curve gradually approaches that of Group C. Groups P 40 C and P are significantly different over Trials 41–60 at the .022 level of confidence while on Trials 101–120 Group P 40 C and Group C do not differ significantly ($P = .447$).

Continuous to partial reinforcement.—Acquisition curves for Groups C 100 P and C 20 P are presented in Fig. 1. In both cases the level of performance drops to that of Group P following the change in reinforcement schedule, the drop being quicker for Group C 100 P. While there is no significant difference ($P = .924$) between Groups C and C 100 P on Trials 81–100 there is a significant difference at the .010 level during Trials 101–121. Trials 101–200 reveal no significant difference between Group C 100 P and Group P ($P = .757$).

The drop in performance for Group C 20 P is more gradual, although a comparison with Group C over the first block of 20 trials on which partial reinforcement was received reveals a significant difference ($P = .015$) and a comparison with Group P over the same trials fails to allow rejection of the null hypothesis ($P = .136$).

An analysis of variance was carried out comparing Groups P, C 100 P, and C 20 P on Trials 101–120. The groups did not differ significantly ($F < 1$).

While Trials 101–111 and 21–30 represent the first blocks of 10 trials during which Groups C 100 P and C 20 P received 50% reinforcement, the first actual nonreinforced trials were not given until trials 102 and 22. Trials 103–112 and 23–32 therefore represent the first 10 trials following the receipt of a nonreinforced trial. The mean percentages of CR's made on these trials were 6.5% and 6.4% lower, respectively, than those made on Trials 101–110 and 21–30. In Fig. 2 the percentage of CR's made by Groups C 100 P and P is plotted by individual trials from Trial 95 to Trial 112. The decrement following the switch to 50% reinforcement appears to occur on the first trial following receipt of a 2400-msec. CS-UCS interval, the percentage of response on that trial (45.0) being little higher than the mean for Trials 111–120 (43.5).

Comparison in terms of number of reinforcements.—Acquisition curves in terms of reinforced trials are presented in Fig. 3. Where partial reinforcement was received, the percentage of response over each block of 10 reinforced trials represents the number of responses made in a 20-trial block divided by the number of responses possible in that block. In the graph Groups C and C 100 P are combined to Trial 100.

When comparisons are made in this manner the conditioning curve of Group P is seen to rise more rapidly. The difference between Groups P and

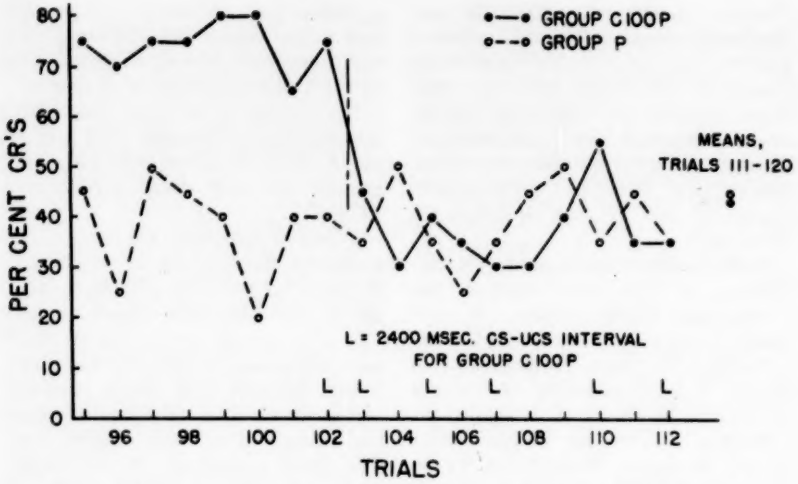

FIG. 2. Mean percentage of CR's by individual trials for Groups C 100 P and P.

C was tested for reinforced Trials 1–20, 21–40, and 91–100. In the latter two cases significant differences were found at the .043 and .021 level, respectively. The null hypothesis cannot be rejected for Trials 1–20 ($P = .070$). The drop of Group C 20 P to the Group P level following its switch to partial reinforcement is more sudden than when the com-

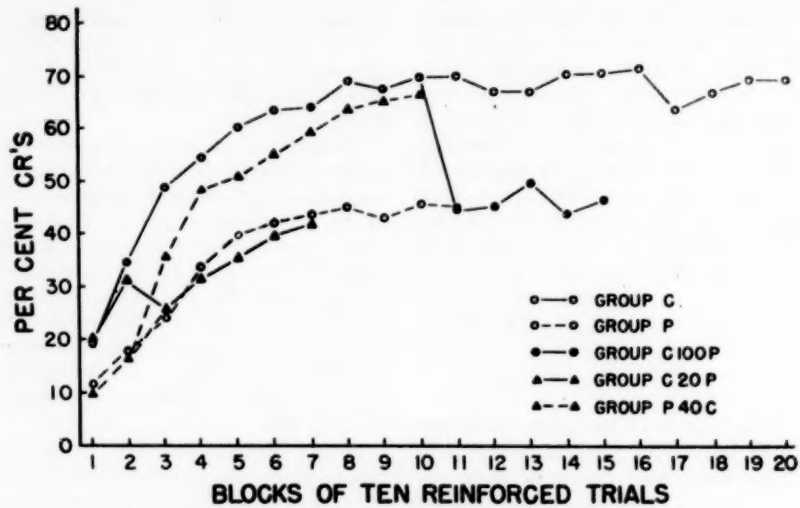

FIG. 3. Mean percentage of CR's in successive blocks of 10 reinforced trials.

parison is made by trials. No significant difference was found between Group C 20 P and Group P over reinforced Trials 21–40 ($P = .667$) or 51–70 ($P = .764$) although after Trial 40 the Group C 20 P curve is consistently below that of Group P.

In an effort to determine the amount of conditioning that took place at the 2400-msec. CS-UCS interval, as well as to see if responses were being made in the 500-msec. period beyond the 200–500-msec. CR interval, trials involving the long CS-UCS interval were scored for the number of responses occurring 500–1000, 1000–1500, and 1900–2400 msec. following the onset of the CS. The results for Group P are presented in Fig. 4. Inspection of the curves reveals that the three intervals were similar both in level of response and amount of increase (approximately 10% over 220 trials). The results for Groups C 100 P and C 20 P were essentially the same, the curves lying between 10% and 30% response with the differences between the means at any point generally less than for Group P. When the interval from 500 to 2400 msec. was scored, the mean percentage of response for Group P was found to be 40.0, 57.5, and 67.5 for Trials 1–20, 81–100, and 201–220.

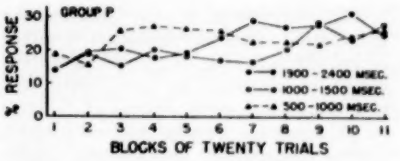

FIG. 4. Mean percentage response in successive blocks of 20 trials made by Group P in the intervals 500–1000, 1000–1500, and 1900–2400 msec. following the onset of the CS.

DISCUSSION

Partial vs. continuous reinforcement.—The results of the present experiment are in agreement with the previous studies which have found higher acquisition performance in classical defense conditioning under conditions of continuous reinforcement. The significant difference between the continuous and partial group curves when compared in terms of reinforced trials indicates the presence of some kind of decremental factor resulting from nonreinforced trials. Since the partial group curve appears to have reached a stable asymptote, the evidence indicates that the extended conditioning period did not result in: (*a*) an increase in performance to the continuous groups level, or (*b*) the decrement in performance that would have been expected if the inhibitory factor had continued to increase after the asymptote of learning had been reached.

Continuous to partial reinforcement.—The sudden drops in performance following switches from continuous to partial reinforcement do not support the notion of a gradual growth of an inhibitory factor over a large number of nonreinforced trials. While a part of the sudden decrement might be attributed to the disrupting effects of a change in procedure per se, it does not seem likely that this could account for the continuing lower performance of the switched groups for the remaining 90–100 trials.

Evidence that the continuing post-shift depression was due to the same factor as that responsible for the partial curve decrement is given by Group C 20 P. In contrast to Group C 100 P, which followed the partial curve's asymptote, this group's conditioning curve rises gradually after its initial decrement, increasing with the growth curve of the partial group. These switch effects, plus the fact that the partial group reached a stable asymptote, suggest that the inhibitory effect of nonreinforced trials reaches a maximum very rapidly, possibly after one nonreinforcement, and remains at a constant value while the partial schedule is continued.

The suddenness of the performance decrement raises the question whether or not the factors acting in this situation might be classified as "set" variables, i.e., related to some degree of "voluntary" control. Several studies (Grant, 1939; Hilgard & Humphreys, 1938; Norris & Grant, 1948) have demonstrated that an inhibitory "set" defined in terms of the instructions given the subject may reduce the frequency of responding. Clarification of the role of such inhibitory set factors, which could result from self-instruction, in the present experiment would seem to be dependent upon further studies concerned with this problem—perhaps involving nonverbal organisms or the use of differential instructions given subjects under different reinforcement conditions.

Partial to continuous reinforcement.—The group switched to continuous reinforcement at Trial 40 gradually approached the performance level of the continuous group. This indicates that dissipation of the inhibitory factor is much slower than its acquisition. It is possible that if more than 20 nonreinforcements were given before the switch the final level of conditioning would remain below that of the partial group. Note should be taken that, in contrast with the groups switched from continuous to partial, the switch from partial to continuous reinforcement did not involve a new type of trial since Group P 40 C had experienced two reinforced trials in succession three times, as well as 14 single reinforced trials, during its partial-reinforcement phase.

The 2400-msec. CS-UCS interval.—In the present study nonreinforcement involved the use of a 2400-msec. CS-UCS interval. In all but one (Reynolds, 1958) of the previous studies cited, nonreinforcement consisted in the presentation of the CS without the UCS being given. The possibility should be considered that, with this technique, performance following a switch to partial reinforcement might be much different. Evidence bearing on this question is provided by a later unpublished study performed in the Iowa laboratory which was primarily concerned with another problem. In this experiment two groups of 28 *S*s each were switched to 50% partial reinforcement for 20 trials following 100 conditioning trials with a 4-lb./sq.-in. air puff UCS. For Group I a nonreinforced trial involved a 2400-msec. CS-UCS interval while for Group II the UCS was omitted with the CS terminating after 500 msec. Both groups dropped significantly from Trials 81–100 to Trials 101-120, Group I dropping 15.36% to a level of 58.03% CR's and Group II dropping 21.9% to a 58.4% CR level. Thus in comparing the two types of nonreinforced trials in switches to 50% reinforcement, omission of the UCS and termination of the CS resulted in a slightly larger decrement in performance than switching to a 2400-msec. CS-UCS interval.

The strong UCS used in this later study resulted in a relatively large number of "voluntary" responders. Analysis of the data for these *S*s revealed a smaller but significant decrement following the switch, the percentage of response dropping 9.17% in the case of *S*s receiving the long interval trials ($N = 12$) and 3.40% for those for whom the UCS was omitted ($N = 12$) as a nonreinforcement condition.

McAllister's finding that a long CS-UCS interval produces little or no conditioning is supported by the analysis of responses made during the long CS-UCS interval trials. Even after 110 reinforcements at 2400 msec. the percentage of response for the 500 msec. preceding the puff is below 30% and not significantly different from the number of responses made 1000–1500 msec. following the onset of the CS. When responses were scored over the whole 500–2400 msec. interval, the increase in percentage of response for Group P from Trials 1-20 to Trials 81-100 was found to be 17.5%, a value that compares closely to the increase (15.5%) found by McAllister when a group that had received a 2500 msec. pairing on each trial was scored

over the interval 100–2500 msec. following the onset of the CS.[2]

Summary

Conditioning of the human eyelid response was carried out for five groups of 20 *Ss* each in order to provide data concerning (*a*) the effects of 50% partial reinforcement during an extended number of acquisition trials, and (*b*) the effects of switches from continuous to 50% and from 50% to continuous reinforcement.

One group of *Ss* received continuous reinforcement for 200 trials while a second group received 50% partial reinforcement for 220 trials. A reinforcement involved a 500-msec. CS-UCS pairing, a nonreinforced trial a 2400-msec. CS-UCS interval. Both groups reached stable asymptotes with the partial group significantly below the 100% group.

Two groups were switched to partial reinforcement, one following 100 continuously reinforced trials and the other after 20 such trials. Totals of 200 and 120 trials, respectively, were given. Both groups dropped quickly to, and remained with, the performance level of the group that had received partial reinforcement throughout training.

The fifth group was given 50% reinforcement for 40 trials and then switched to continuous reinforcement, which was continued for 80 trials. Following the switch the conditioning curve gradually approached that of the continuous group, being not significantly different from it during the last block of 20 trials.

The significant difference between the 50% and 100% reinforcement groups after an extended conditioning period was interpreted as indicating the presence of some kind of decremental factor resulting from nonreinforced trials. Since the 50% group reached a stable asymptote, neither a continued growth of an inhibitory factor nor a continued increase in performance to the 100% level was indicated. The sudden decrements in performance following switches to 50% reinforcement indicates that the inhibitory effects of nonreinforced trials reach a maximum very quickly, possibly after one nonreinforcement. The more gradual approach of the group switched from 50% to 100% reinforcement to the 100% group's performance level suggests that dissipation of the inhibitory factor is much slower than its growth with nonreinforcements. Analysis of the responses made during the long CS-UCS interval trials supports McAllister's finding that such an interval produces little or no conditioning.

[2] Myers (1950) found a 10% increase in random blink rate when a 3000-msec. visual CS was given alone for 80 trials and a 12.7% increase when a group was given 80 presentations of an air puff alone. In both cases the interval measured was 2300 msec.

References

Grant, D. A. The influence of attitude on the conditioned eyelid response. *J. exp. Psychol.*, 1939, **25**, 333–346.

Grant, D. A., & Hake, H. W. Dark adaptation and the Humphreys random reinforcement phenomenon in human eyelid conditioning. *J. exp. Psychol.*, 1951, **42**, 417–423.

Grant, D. A., Riopelle, A. J., & Hake, H. W. Resistance to extinction and the pattern of reinforcement. I. Alternation of reinforcement and the conditioned eyelid response. *J. exp. Psychol.*, 1950, **40**, 53–60.

Grant, D. A., & Schipper, L. M. The acquisition and extinction of conditioned eyelid responses as a function of the percentage of fixed-ratio random reinforcement. *J. exp. Psychol.*, 1952, **43**, 313–320.

Grant, D. A., Schipper, L. M., & Ross, B. M. Effect of intertrial interval during acquisition on extinction of the conditioned eyelid response following partial reinforcement. *J. exp. Psychol.*, 1952, **44**, 203–210.

Gynther, M. D. Differential eyelid conditioning as a function of stimulus similarity and strength of response to the CS. *J. exp. Psychol.*, 1957, **53**, 408–416.

Hilgard, E. R., & Humphreys, L. G. The effect of supporting and antagonistic voluntary instructions on conditioned discrimination. *J. exp. Psychol.*, 1938, **22**, 291–304.

Humphreys, L. G. The effect of random alternation of reinforcement on the acquisition and extinction of conditioned eyelid reactions. *J. exp. Psychol.*, 1939, **25**, 141–158.

Hunter, T. A., & Brown, J. S. A decade-type electronic interval timer. *Amer. J. Psychol.*, 1949, **62**, 570–575.

Kimble, G. A., Mann, L. I., & Dufort, R. H. Classical and instrumental eyelid conditioning. *J. exp. Psychol.*, 1955, **49**, 407–417.

McAllister, W. R. The effect on eyelid conditioning of shifting the CS-US interval. *J. exp. Psychol.*, 1953, **45**, 423–428. (a)

McAllister, W. R. Eyelid conditioning as a function of the CS-US interval. *J. exp. Psychol.*, 1953, **45**, 417–422. (b)

Myers, J. A. An experimental investigation of the effect of varying the time between the onsets of the conditioned and unconditioned stimuli on the conditioned eyelid response. Unpublished Ph.D. dissertation, State Univer. of Iowa, 1950.

Norris, E. B., & Grant, D. A. Eyelid conditioning as affected by verbally induced inhibitory set and counter re-enforcement. *Amer. J. Psychol.*, 1948, **61**, 37–49.

Reynolds, W. F. Acquisition and extinction of the conditioned eyelid response following partial and continuous reinforcement. *J. exp. Psychol.*, 1958, **55**, 335–341.

Spence, K. W., & Taylor, J. A. Anxiety and strength of the UCS as determiners of the amount of eyelid conditioning. *J. exp. Psychol.*, 1951, **42**, 183–188.

(Received March 17, 1958)

Journal of Experimental Psychology
Vol. 57, No. 2, 1959

THE DIFFERENCE THRESHOLD OF THE MAGNITUDE OF VISUAL VELOCITY [1]

BOB B. BRANDALISE [2] AND ROBERT M. GOTTSDANKER

University of California, Santa Barbara

Despite an early start in experimentation, there have been few studies of the difference threshold of the magnitude of visual velocity, and these have not yielded definitive values. Aubert's (1886) initial investigation of the problem indeed was so unsystematic that it defies attempt at summarization. There have been two subsequent studies. Bourdon (1902, p. 192) reported threshold values based upon data from only one *S*. In the investigation by Brown and Mize (1932) there were only two *S*s, each of whom made but a small number of judgments at any one standard. Further, in both cases there was investigated only a limited range of rather slow motions.[3]

The present study was designed to make a new assessment of the difference threshold under conditions which eliminate the shortcomings of earlier work and which insure accurate measurement. A satisfactory determination would fill an obvious gap in our psychophysical knowledge of perceptual abilities, and might provide a basis for expectations in tasks requiring the reproduction or matching of motions in the environment, e.g., tracking.

METHOD

The *S* observed two rotating black disks. Each had a small white dot near its edge. His task was to compare and to equate the velocity of the two target dots.

Apparatus

Display.—The *S* was seated 2 m. from the target display. The disks were positioned perpendicular to his line of sight on a flat white background panel. Following are the pertinent dimensions: diameter of disks, 100 mm.; diameter of target dots, 12 mm.; distance of center of dot from center of disk, 90 mm.; distance between centers of disks, 42 cm.

A partition was placed between the two disks so that *S* could see only one at a time when his chin was in the rest mounted on either side of the partition. A separate spotlight was used for each disk. This furnished an illumination of 48 ft.-c. as measured with a Macbeth Illuminometer.

The disk on *S*'s left provided the standard motion. Its rotational velocity was held constant for a block of trials. The velocity of the other disk was adjustable by *S* through a continuously variable and unmarked control wheel mounted on a black box near his right hand.

Standard-motion units.—Five synchronous "Circle B" motors with the disks attached directly to their shafts produced the five

[1] This article is based on the first writer's masters thesis at the University of California, Santa Barbara. The research was supported by the U. S. Air Force under contract AF 33(616)-2024 with the University of California monitored by the Psychology Branch of the Aero Medical Laboratory, Directorate of Research, Wright Air Development Center. Permission is granted for reproduction, publication, use, and disposal in whole or in part by or for the U. S. Government.

Appreciation is expressed to Robert W. Reynolds for his valuable and time-consuming contributions to the project.

[2] Presently at Sandia Corporation, Albuquerque, New Mexico.

[3] Mention should be made of the more recent study in which Hick (1950) found threshold values for sudden changes in velocity, over a wide range of standard velocities. Although *S* has a far different task than when comparing two ongoing motions, it is of interest that his threshold values are quite comparable to whose that have been obtained for the latter task.

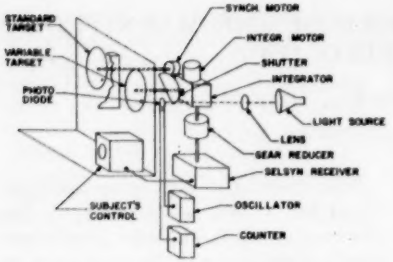

FIG. 1. Schematic of the apparatus.

standard magnitudes of velocity. The rotational rates were 10, 20, 30, 60, and 90 rpm. In Table 1, which presents the results, the velocities are also specified in two other ways. The first is the *path* velocity of the target dot in millimeters per second. The second is the *angular displacement* velocity at the eye of the *S* in degrees of visual angle per second.

During the experiment the standard-motion units were interchanged for each block of trials.

Variable-motion unit.—The *S* adjusted the rotational rate of the variable target by his remote control positioning of the speed control on a precision ball-and-disk integrator. Output of the integrator could be varied continuously from a rotational rate of zero to three times that of the constant input. Three synchronous Bodine motors, rated at 15, 25, and 60 rpm, were interchanged as the constant input to the integrator according to the standard-motion unit employed.

The *S*'s control wheel was fixed to the rotor of a selsyn transmitter. The rotor of the selsyn repeater fed through a 30:1 reduction gear box to the speed control of the integrator. The entire range of output rates for any one constant input motor was covered in 48 complete rotations of *S*'s control wheel. Figure 1 shows a schematic of the experimental apparatus.

Timing circuit.—Rotational rates of the target disks were measured by an electronic modification of the mechanical method reported by Klein (1942). The basic circuit consisted of an audio oscillator, a photo diode, and an electronic counter. A 180° shutter mounted on the output shaft of the integrator (or on one of the standard-motion units when it was being calibrated) was located between the photo diode, which was mounted on the reverse side of *S*'s background panel, and a spotlight 2 m. away. Light was focused through a lens so that a narrow beam of light fell on photo diode through 180° of shaft rotation. Stability of performance was improved by a blue filter which reduced the infrared rays.

When the photo diode was activated, its resistance decreased, shorting out the oscillator input to the counter. During the time that the 180° shutter shaded the diode, the counter received 700 impulses per sec. from the oscillator. Rotational rate of the target disk thus could be computed. Frequency output of the oscillator varied only .21% after 5 hr. of continuous operation.

Calibration.—For each of the five standard units, the number of impulses in 180° of rotation was determined on 100 measurements. Speeds based on the mean counts deviated between .09% and 1.37% from the rated speeds of the five motors. The maximum relative *SD* around any mean was .46% of the mean value.

When the method of measuring performance on the variable-motion was calibrated, the shutter was attached to the output shaft of the integrator, where it remained for the rest of the experiment. Adjustment was made of the variable target's speed until it seemed to be the same as the slowest standard speed, and 100 readings were taken on the counter, with *S*'s control wheel stationary. This process was repeated for the other four standard speeds. Using the average of three readings on each adjustment (as was done in the experiment proper), the maximum relative *SD* attributable to this method of measurement is .26%.

Since the lowest relative threshold (and a very atypical one) reported previously was 2.4% of the standard velocity (Brown & Mize, 1932), it was felt that the apparatus and techniques of measurement were entirely satisfactory for present purposes.

Subjects

Five men and five women, undergraduate majors in psychology at the University of California, Santa Barbara, were used as *S*s. All had 20/20 vision (Snellen) and were naive regarding the task.

Procedure

The psychophysical method of average error (or adjustment) was used. The *S* was permitted to take as long as he wished in making each adjustment and was allowed to bracket his responses. No restrictions were placed on the manner in which judg-

TABLE 1

SD's (Thresholds) and Constant Errors[a]

S	Rotational (RPM), Path (Cm./Sec.), and Angular Displacement (Deg./Sec.) Velocities of Standard									
	10 RPM 9.4 Cm./Sec. 2.7°/Sec.		20 RPM 18.8 Cm./Sec. 5.4°/Sec.		30 RPM 28.3 Cm./Sec. 8.1°/Sec.		60 RPM 56.5 Cm./Sec. 16.2°/Sec.		90 RPM 84.8 Cm./Sec. 24.3°/Sec.	
	SD	CE	*SD*	CE	*SD*	CE	*SD*	CE	*SD*	CE
1	6.00	−0.33	5.13	−0.33	8.09	−4.72	4.51	0.73	4.04	0.13
2	9.96	3.85	8.96	0.37	8.48	1.59	5.47	3.18	6.31	−1.72
3	7.56	3.26	5.33	−2.44	4.27	−2.09	3.04	1.24	5.93	−3.48
4	7.52	8.82	5.52	9.67	6.16	7.25	6.57	5.69	7.65	2.06
5	15.59	−1.18	11.09	1.18	9.80	0.46	6.88	5.10	6.67	−0.69
6	11.37	1.22	9.52	2.11	9.78	−2.18	9.73	1.57	6.79	−1.42
7	7.37	0.59	5.35	0.22	5.14	0.67	4.99	1.21	7.17	−5.16
8	9.22	2.41	7.94	−1.24	8.69	−5.94	9.16	−0.96	7.73	−5.07
9	14.30	9.96	7.72	2.98	9.67	0.24	6.40	8.23	6.36	3.09
10	5.96	−0.97	3.35	1.30	1.93	−1.28	1.98	1.24	5.54	−7.90
Mean	9.49	2.78	6.99	1.26	7.20	−0.60	5.87	2.72	6.42	−2.02

[a] Percentages of magnitudes of standard velocities.

ments were made: *S* was encouraged to use whatever method suited him best.

Typically, *S* took about a minute in all to make an adjustment, looking at the standard and the variable each two or three times. No one was observed to look back and forth in an attempt to synchronize the two dots, nor is it likely that this could have been done because of the chance relation in positions of the dots.

Several of the *S*s reported counting elapsed time to cover a zone in the case of the lowest magnitude. In all instances this method was said to have been dropped very soon. When asked whether rotational tempo (as tapping once per revolution) had been employed, all *S*s reported negatively. It seems that the very large number of trials discouraged adjusting on any basis other than general appearance, since the more analytical techniques required a much greater expenditure of effort.

Each *S* underwent two 3-hr. test sessions, spaced a day apart. In the first session he started on the lowest standard magnitude of the range and worked up through the highest, making 25 adjustments at each standard. In the second session he worked in the reverse order. Thus *S* had 50 perceptual matches to make at each standard magnitude of velocity. A preselected sequence including 25 values clearly above the standard and 25 values clearly below was used by *E* in setting the initial velocity of the variable disk on the 50 trials. A 10-min. rest period was introduced halfway through each session.

Results

Table 1 shows the relative *SD*'s and constant errors of the 10 *S*s at each of the five standard magnitudes of velocity. Values are given as percentages of the standard magnitude in question.

Constant errors in many cases are sizable. Obvious sources of bias are position, sequence, and range effects. However, the proper controls for the study of these effects were not imposed as they would greatly have complicated the testing procedure and analysis of data.

The *SD*'s provide the threshold measures which were sought. It may be noted that the mean *SD*'s range from 5.87% for the next-to-highest standard rate to 9.49% for the lowest, with an over-all mean of about 7%. There is no evidence for a U-shaped function.

The analysis of variance in Table 2 shows both Magnitudes and *S*s to be significant at the .01 level. How-

TABLE 2

ANALYSIS OF VARIANCE OF THRESHOLDS

Source	*df*	*MS*	*F*	*P*
Magnitudes	4	19.10	7.39	.01
Ss	9	21.20	8.21	.01
M × *S*	36	2.58		

ever, only comparisons involving the lowest magnitude provide critical differences at the .01 level, or indeed at the .05 level, the required differences being 1.96% and 1.46%, respectively (Lindquist, 1956, p. 166). Further, when the lowest magnitude was removed from the analysis of variance, the mean square associated with magnitudes was reduced from 19.10 to 3.28, which in turn gave an insignificant *F* of 1.44. The large and consistent variation among *Ss* is noteworthy. While S_5 and S_8 never descended below a threshold of 6.6%, S_{10} was below 2% on two of the speeds.

DISCUSSION

It appears that Weber's law of the invariance of relative threshold is applicable for all but the lowest magnitude of velocity tested, since significant differences among the other magnitudes could not be demonstrated, even at a low level of confidence. It is surprising that the threshold remained low for the highest magnitude since some *Ss* complained that the target was "blurry" and hard to estimate.

A comparison may be made between the values found in this study with those obtained by Bourdon (1902) and those obtained by Brown and Mize (1932). In all three investigations, the same viewing distance, 2 m., was employed. However, as is pointed out by Brown and Mize, it is not certain that the same values should be expected to hold for translatory motion, such as they used, and for rotary motion, such as Bourdon (and the present study) employed. There remains the possibility that the values depend upon the rotary rate of the disk rather than the distance moved by the target. Information on this problem would be provided by a study in which the pair of disks was placed at different distances from *S* and in which the distance of the dot from the center of the disk was varied. Finally, there is always a difficulty in drawing firm conclusions from values derived by different psychophysical procedures.

Some adjustment is necessary for the values Bourdon reported with the method of limits. Each of his increments was an increase of 1/50 of the standard speed. It seems likely that his threshold values of 1/12 and 1/8 refer to the fourth and sixth increments, respectively. Therefore, in the first case the .50 probability threshold would fall between the third and fourth steps, and in the second case between the fifth and sixth steps. The best estimates of threshold by the .50 probability criterion are 7% and 11%. The threshold values reported by Brown and Mize are already in the form of .50 probabilities.

When the method of average error is employed, the meaning of .50 probability of detection is less clear. Probably the best estimate is the point which marks off the probable error, half of the adjustments being within this range and half without. Present relative thresholds in terms of *SD*'s were converted to probable errors by multiplying by .6745.

Findings of the three investigations on difference thresholds are compared in Table 3. The most striking fact is that present values are consistently the lowest, being generally below 5% as compared with the other values, which tend to run over 10%. It is difficult to see how very much of this effect can be attributed to the somewhat different concept of a .50 probability of detection with the method of average error. Probably more important is the difference in the nature of a "trial"

in this method. First, *S* had unlimited time to observe the motion. It may be suggested in the method of limits that *S* feels obliged to make a choice at each step within a fairly short time. Second, in the present method *S* could have several effective trials in one recorded trial as he was free to bracket any adjustment he had made without having it count. This would naturally reduce the variability of his choices. One other point is that Bourdon's threshold values would perhaps have been lower had he not maintained a point of fixation in the center of the disk.

There are some other points of interest in the comparison. The extreme dip in relative threshold reported by Brown and Mize for their 10 cm./sec. (2.9°/sec.) motion wasn't found in the present study, nor is there a comparable dip anywhere else in the table. Within the range of magnitudes employed, the present study shows more consistency in general than has been found previously.

The present study is the first to give any idea of the range of variation between *Ss* in this threshold. Because no obvious differences in technique of viewing were observed or reported it might have been expected that such variation would be moderate among *Ss* of similar age and background. On the contrary, cases were found for some velocities in which the value for one *S* was five times that for another *S!* A number of possible explanations suggest themselves. High values might be due to inattention, fatigue, or confusion. Again, there may have been discrepancies in how the motion was seen (e.g., as rotary motion of a disk or as motion of a dot along a circular path). It may be that some *Ss* had learned how to evaluate rates of motion. Unfortunately, the present study can only pose the question, not provide the solution.

Summary

Ten adult *Ss* were given the task of adjusting the rotational rate of one disk so that a small target dot near its edge appeared to be moving at the same speed as the target dot on another rotating disk. Rotational rate of the comparison disk was set at one of five values between 10 and 90 rpm. These provided path velocities of the target dot between 9.4 and 84.8 cm./sec., and angular displacement velocities between 2.7 and 24.3°/sec. of visual angle. Thresholds were defined by the *SD* of adjustments of the ten *Ss*.

TABLE 3

Relative Difference Thresholds (.50 Probability)

Study	Standard Magnitude of Velocity			Difference Threshold (%)
	Rotational RPM	Path (Cm./Sec.)	Ang. Displacement (Deg./Sec.)	
Present:	10	9.4	2.7	6.40
Circular target motion;	20	18.8	5.4	4.71
Method of average error	30	28.3	8.1	4.86
	60	56.5	16.2	3.96
	90	84.8	24.3	4.33
Bourdon (1902):	2.3	2.7	0.8	11.0
Circular target motion;	7.1	8.3	2.3	7.0
Method of limits	15.0	17.6	5.0	7.0
Brown and Mize (1932):	—	6	1.7	14.5
Linear target motion;	—	8	2.2	12.6
Method of limits	—	10	2.9	2.4
	—	12	3.4	11.2
	—	14	4.0	16.9
	—	16	4.6	7.4

Mean values obtained relative to that of the standard magnitude which was being matched ranged from 5.87% to 9.49%. Only the slowest speed tested differed statistically from the others, it giving rise to the highest value obtained.

Present findings were compared with those of other investigators. All thresholds were expressed as the relative difference required for a .50 probability of discrimination. For the present study this was accomplished by multiplying the mean *SD* by .6745. Values obtained here tended to be slightly below 5%, or about half the size of those found previously. An attempt was made to relate discrepancies in results to differences in methods of experimentation. Note was also made of the marked variability among *S*s in the present study.

References

AUBERT, H. Die Bewegungsempfindung. *Arch. f. d. ges. Physiol.*, 1886, **39**, 347–370.

BOURDON, B. *La perception visuelle de l'espace.* Paris: Schleicher Freres, 1902.

BROWN, J. F., & MIZE, R. H. On the effect of field structure on differential sensitivity. *Psychol. Forsch.*, 1932, **16**, 355–372.

HICK, W. E. The threshold for sudden changes in the velocity of a seen object. *Quart. J. exp. Psychol.*, 1950, **2**, 33–41.

KLEIN, G. S. The relation between motion and form acuity in para-foveal and peripheral vision and related phenomena. *Arch. Psychol., N. Y.*, 1942, No. 275.

LINDQUIST, E. F. *Design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin, 1953.

(Received February 7, 1958)

Journal of Experimental Psychology
Vol. 57, No. 2, 1959

INFORMATION AND AUTOKINETIC MOVEMENT [1]

WARD EDWARDS

University of Michigan

The ordinary visual world looks stable and stationary. Often, however, *S*s will report that an objectively stationary dot or pattern of light in an otherwise dark room is seen as moving; this illusion is called *autokinetic movement*. What properties determine whether or not a stimulus will appear to move autokinetically? Or, to put the question another way, why do not ordinary visual stimuli, like this page as you read it, seem to move?

The usual stimulus for autokinetic movement is dim and small. But two previous experiments (Edwards, 1954a, 1954b) show that dim stimuli subtending 60° or more of visual angle and bright stimuli (174 ft.-L.) subtending 7½° regularly produce autokinetic movement. These are not upper limits; presumably even larger and brighter stimuli would produce autokinetic movement if prevented from illuminating other objects. Clearly, neither the large size nor the brightness of this page is what stabilizes it.

One of the previous experiments (Edwards, 1954a) showed that a rectangular matrix of dots stimulated more autokinetic movement than the same matrix with all dots masked except the outer rectangle and the central fixation point. This finding violates the principle derived from many previous experiments that the more stimuli present, the less autokinetic movement will be seen. It therefore suggests that autokinetic movement is associated with a monotonous visual world, one which either has very few objects in it, or else has many objects but arranged in a regular and meaningless way. Perhaps autokinetic movement occurs "with a highly redundant visual world and does not occur when the visual world is full of information. This principle . . . would suggest that an irregularly arranged field of dots should show less movement than a regular one containing the same number of dots in about the same total extent" (Edwards, 1954a, p. 495). The present experiment examines that prediction.

Method

Apparatus.—A large light box into which masks could be inserted, plus appropriate timers and switches, was the apparatus. A mirror system extended the distance light traveled from box to *S* to 16.4 ft. The stimulus patterns subtended 11°. There were two patterns. One was a square matrix, 17 holes on a side. Each hole was ⅛ in. in diameter; the distance between centers of adjacent holes was 1¼ in. The central hole was the fixation point; it was smaller, brighter, and whiter than the others. The other pattern was also square in shape, but its holes were randomly arranged. The four corners and the central fixation point were identical in both patterns. The number of holes on the lines connecting the corners was 17 in both patterns, but the spacing was irregular for the random pattern. The interior holes of the random pattern were randomly arranged, with overlapping holes forbidden. The total number of holes was 289 for both patterns.

Subjects.—The *S*s were 20 basic airmen at

[1] This research was done at the Systems Operator Research Unit, Operator Laboratory, Air Force Personnel and Training Research Center, under ARDC Project 7737, Task No. 27107, in accordance with ARDC Technical Requirement No. 181-56, in support of the research and development program of the Air Force Personnel and Training Research Center. Permission is granted for reproduction, translation, publication, use and disposal in whole or in part by or for the United States Government. The author is grateful to James L. McDade, who ran *S*s and analyzed the data.

Lackland Air Force Base. All were unacquainted with autokinetic movement under laboratory conditions.

Procedure.—Each *S* entered the experimental room in the dark. During 10 min. of dark adaptation, *E* told *S* that he would be shown patterns of dots which would start to move after a few seconds. When movement began, *S* pressed a button. This started a timer which recorded latency and another which turned the stimulus off after 30 sec. Whenever the stimulus was on, a motor which served as a sound effect was also on. Each mask was used 10 times. The order of presentation was counterbalanced.

Results

The results of this experiment disprove the hypothesis about information and autokinetic movement on which the experimental design was based. All *S*s saw autokinetic movement on all trials. Latencies decreased and amounts of movement seen increased for each *S* as the experiment progressed. The pattern of results was essentially the same as that found in a previous experiment (Edwards, 1954a), and so needs no discussion here. But there were no significant or even suggestive differences between the random and the regular masks in amounts or latencies of movement seen.

If this experiment is a fair test, then the amount of information in a stimulus does not control the amount or latency of autokinetic movement which that stimulus induces. This still leaves unanswered the question: Why do not ordinary visual stimuli produce autokinetic movement? An attractive next hypothesis to try is the following: Autokinetic movement of complex patterns is mostly movement of the whole pattern, rather than movement of parts with respect to other parts. This means that if contours are located near the edge of the visual field, sustained movement in a single direction would require that these contours vanish from the visual field—which they cannot do so long as *S* does not move his eyes or head. Ordinary visual experience is full of contours near the edge of the visual field; perhaps they are what keep ordinary stimuli stable.

Summary

This experiment examines the hypothesis that autokinetic movement occurs if the visual world is highly redundant and does not occur if the visual world is full of information. Specifically, it measures the amount of autokinetic movement produced by two stimuli: an orderly and an equivalent disorderly array of dots. The two arrays produced equal amounts and latencies of movement. If this experiment is a fair test, then the hypothesis is wrong.

References

Edwards, W. Autokinetic movement of very large stimuli. *J. exp. Psychol.*, 1954, **48,** 493–495. (a)

Edwards, W. Two- and three-dimensional autokinetic movement as a function of size and brightness of stimuli. *J. exp. Psychol.*, 1954, **48,** 391–398. (b)

(Received March 3, 1958)

Journal of Experimental Psychology
Vol. 57, No. 2, 1959

FINGER SPAN: RATIO SCALE, CATEGORY SCALE, AND JND SCALE [1]

S. S. STEVENS AND GERALDINE STONE

Harvard University

This is an account of the experiments on finger span referred to in an earlier paper (S. S. Stevens, 1957). The purpose was to determine, by the method of magnitude estimation, how the subjective impression of thickness varies with the width of objects grasped between the thumb and the middle finger, and to compare this function with both the category scale and the Fechnerian scale constructed from just noticeable differences (JND's). The problem is whether finger span behaves as a prothetic continuum, whether the ratio scale of subjective thickness grows as a power function of stimulus width, and whether the category and JND scales exhibit the same relations to the ratio scale that characterize other prothetic continua.

Magnitude Estimation

The scaling of finger span was begun in 1956 but was interrupted by a fire that consumed the tower of Memorial Hall and resulted in piles of soaked stimuli and data sheets. The experiments were finally finished in 1958. A total of five experiments were conducted under various modifications of the method of magnitude estimation, but since they all gave similar results only two experiments will be described here.

[1] Supported by a grant from the National Science Foundation and by a contract with the Office of Naval Research (Project NR142-201, Report PNR-214). Reproduction for any purpose of the US Government is permitted.

Method.—Sets of wooden blocks, about 6 in. long, were cut from $\frac{3}{4}$-in. pine. One set of 13 blocks varied in width from 2.3 to 63.7 mm. in roughly equal steps. Another set of 12 blocks varied from 5 to 65 mm. in steps of 5 mm. The first set was used in the two experiments to be described.

The *O* (eyes closed) rested his wrist on a padded arm rest and held his thumb and middle finger about 80 mm. apart. The *E* brought one of the blocks into contact with *O*'s thumb, whereupon *O* brought his middle finger into contact with the other side of the block and then made his judgment. Typically, a standard block was presented first and *O* was told to call it 10, or some other value. He was then asked to assign numbers to the other blocks proportional to the apparent thickness. In some experiments the standard was presented only once at the beginning of the run. In other experiments the standard was presented before each variable.

In each experiment 10 *O*s made two judgments of each stimulus. The order of presentation was irregular and different for each *O*. The *O*s were staff, students, and visitors of various sorts, a total of 33 different people.

Results.—The geometric means of the magnitude estimates are plotted in Fig. 1. For the upper curve (triangles) a standard of 32.4 mm. was presented only once at the beginning of the experiment and *O* ($N = 10$) was told to call it 100. For the lower curve (circles) the standard was 13.7 mm. (called 10) and was presented ($N = 10$) before each variable. Except for two people, different *O*s were used in the two experiments.

The fact that the data fall close to a straight line in a log-log plot shows that apparent thickness is a power function of stimulus width. The slope of the line, which gives the exponent of the power function, is

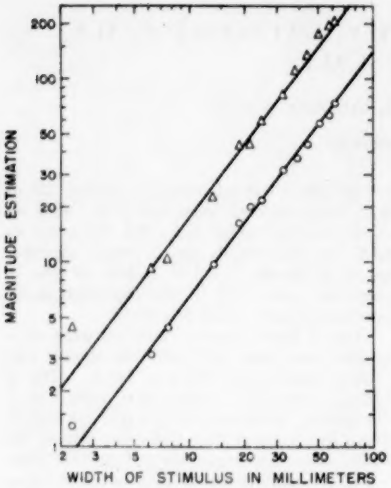

FIG. 1. Magnitude estimations of finger span plotted in log-log coordinates. *Triangles:* a standard of 32.4 mm. was presented at the beginning of the experiment and called 100. *Circles:* a standard of 13.7 mm. (called 10) was presented before each stimulus to be judged. Points are geometric means of 20 estimates, two by each of 10 *Os*.

1.33. In other words, apparent thickness is a positively accelerated function of stimulus magnitude.

The lower curve in Fig. 1 may be used to define a scale of subjective finger span. The unit can be taken as the apparent thickness of a stimulus 2.5 mm. thick. This unit will be called a *pak*, which is derived (with some license) from the Greek word for thickness. In terms of this unit, the scale defined by the lower curve in Fig. 1 is described by

$$P = .296S^{1.83}$$

where P is subjective thickness in paks and S is stimulus width in millimeters.

This equation seems to hold very well for stimuli thicker than about 5 mm. On the other hand, as is shown by the bottom points on both curves in Fig. 1, there is some suggestion that for very thin stimuli the formula may break down.

Although only the geometric means of the judgments are plotted in Fig. 1, it should be noted that there was no systematic difference between these geometric means and the medians of the judgments. It is also interesting that the extreme range of the estimates made for a given stimulus was seldom greater than 4 to 1. This is a smaller range of variability than is usually found with such continua as loudness (S. S. Stevens, 1956) and electric shock (Stevens, Carton, & Shickman, 1958).

CATEGORY SCALE

Extensive evidence from other perceptual continua (Stevens & Galanter, 1957) suggests that if finger span is a prothetic continuum, the category scale should be nonlinearly related to the scale obtained by magnitude estimation. This expectation was borne out in two separate experiments, one of which will be described.

Method.—Except for one crucial difference the procedure was identical to that used for magnitude estimation. The stimuli were the same as those for the experiments in Fig. 1. Instead of assigning a number to only one stimulus and leaving *O* free to use any

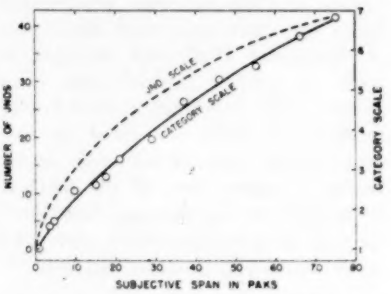

FIG. 2. Category scale and JND scale for finger span plotted against the subjective scale determined by the lower line in Fig. 1. Circles are means of 20 judgments, two by each of 10 *Os*. The JND scale was determined from the function in Fig. 3.

numbers he chose, *E* presented a stimulus of 2.3 mm., which *O* was to call 1, and a stimulus of 63.7 mm., which *O* was to call 7. The *O* was then instructed to partition the continuum into categories on a scale from 1 to 7. Each stimulus was presented twice to each of 10 *O*s in an irregular order, a different order for each *O*.

Results.—The mean category assignments are shown by the circles in Fig. 2, where the abscissa is the scale of subjective finger span in paks. This scale is taken directly from the lower curve in Fig. 1. Plotted against the scale of subjective magnitude, the category scale is concave downward.

The reason for the curvature of this category scale is presumably the same as the reason for the curvature of other such scales (see Stevens & Galanter, 1957). On prothetic continua there is a basic asymmetry in a person's sensitivity to differences as he goes from the low to the high end of the continuum. A given difference is more impressive at the low end than at the high end, with the result that *O* makes his categories a little smaller at the low end and a little larger at the high end. On a continuum such as finger span, which is relatively easy to judge, the curvature is not very great, but it is definitely present. It is comparable to the curvature found in category judgments of visual length (Stevens & Galanter, 1957).

Just Noticeable Differences

Although the size of the JND for finger span has been reported for single, isolated standards (Kelvin, 1954; Langfeld, 1917), there appear to be no published results on the size of the JND as a function of stimulus magnitude. Fortunately, an extensive study was carried out in 1934 by G. J. Huberman (under the direction of J. Volkmann) and the results are contained in an Honors Thesis still on file at Harvard University.

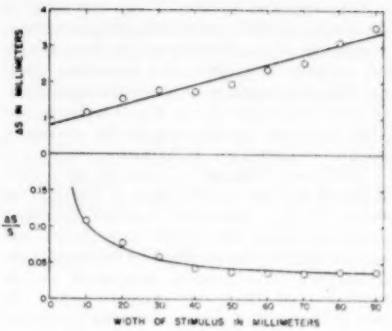

Fig. 3. Differential sensitivity for finger span.

The method of single stimuli, involving two stimuli and two categories, was used. The *O* judged each stimulus as either "long" or "short," and the stimuli were chosen in such a way that the larger stimulus was judged "long" three or four times as often as the smaller stimulus. In this way a stable estimate was obtained of the *SD* of the psychometric function. This value is taken as the JND, or ΔS.

The stimuli consisted of carefully machined blocks of fiber mounted on the rims of wheels in such a way that they could be quickly and easily presented for judgment. Three well-trained *O*s completed the entire series, making a judgment every 8 sec. The average values of ΔS for these *O*s are shown in Fig. 3.

The upper plot shows that ΔS is approximately a linear function of the size of the stimulus. The equation for the line through the data is

$$\Delta S = .0286S + .8$$

As suggested elsewhere (S. S. Stevens, 1957) a linear equation of this general form applies to a wide variety of prothetic continua. It appears to be the general rule to which there are few if any exceptions. It is, of course, the familiar Weber's law modified by the addition of a constant.

The lower plot in Fig. 3 shows the same data plotted in terms of the Weber fraction ΔS/S. The familiar decreasing function is evident in this plot. The curve in the lower plot corresponds to the straight line in the upper plot.

The function in Fig. 3 can be used to determine a JND scale for finger span. Such a scale can be constructed simply by counting off successive JND's and recording their number as a function of stimulus magnitude. Or since the upper plot in Fig. 3 is linear, the JND scale can be obtained by the process of integration (see Luce & Edwards, 1958).

The scale obtained by counting off JND's is shown by the dashed curve in Fig. 2. As seems to be universally true on prothetic continua, when the JND scale is plotted against the ratio scale obtained by magnitude estimation, it is concave downward and is more curved than the category scale. It should be noted, however, that the curvature of category scales can be modified by stimulus spacing, and can, under some circumstances, be made to approximate the curvature of the JND scale (Stevens & Galanter, 1957).

Discussion

At the outset of these experiments it was not self-evident that finger span would satisfy the criteria (S. S. Stevens, 1957) of a prothetic continuum. Since *O* may in effect judge the *position* of his fingers, there was a reasonable possibility that this continuum would turn out to resemble the metathetic continua, such as pitch, visual position (azimuth), and visual inclination. On these metathetic continua *O*'s sensitivity to differences, measured in subjective units, remains approximately constant from one end of the continuum to the other. The curvature of the dashed line in Fig. 2 shows, however, that finger span behaves as a prothetic continuum on which the subjective size of the JND increases as one goes up the scale. Thus it appears that the judgment of finger span is more a judgment of length than a judgment of position.

As with the other prothetic continua, the three procedures (ratio scaling, category scaling, and JND scaling) result in three different functions. Obviously three different aspects of finger span are being measured by these three functions. Although a certain amount of argument has revolved around the question of which of these functions is the "true" scale, it should be apparent that all three are true scales of something or other. All three may have their uses in particular circumstances and the choice among them ought properly to depend on the problem at hand. The JND scale answers questions about resolving power; the ratio scale obtained by magnitude estimation tells how apparent thickness depends on stimulus thickness; and the category scale tells how *O* partitions a segment of the continuum. The category scale is probably less useful, in general, than either of the other two scales, mainly because its form is critically dependent on the spacing of the stimuli used to obtain the scale. The form of the ratio scale is relatively independent of stimulus spacing (J. C. Stevens, 1958; S. S. Stevens, 1956). It is this ratio scale that seems most directly to answer the original problem of psychophysics as formulated by Fechner—namely, How does the magnitude of the subjective response vary with stimulus input?

Summary

A scale for the apparent thickness of stimuli held between the thumb and middle finger was constructed by the method of magnitude estimation. Apparent thickness grows as the 1.33 power of stimulus magnitude. Judgments of thickness on a 7-point category scale yielded a function that is concave downward when plotted against the ratio scale obtained by magnitude estimation.

The difference limen ΔS grows as a linear function of stimulus thickness S. The scale obtained by counting off JND's was plotted against the ratio scale obtained by magnitude estimation. The result is a curve, concave downward, that is more curved than the category scale.

The relations among the three scales (ratio, category, and JND) suggest that finger span is a prothetic continuum and that the JND increases in subjective size as the stimulus is increased.

References

Kelvin, R. P. Discrimination of size by sight and touch. *Quart. J. exp. Psychol.*, 1954, **6**, 23–34.

LANGFELD, H. S. The differential spatial limen for finger span. *J. exp. Psychol.*, 1917, **2**, 416–430.

LUCE, R. D., & EDWARDS, W. The derivation of subjective scales from just noticeable differences. *Psychol. Rev.*, 1958, **65**, 222–237.

STEVENS, J. C. Stimulus spacing and the judgment of loudness. *J. exp. Psychol.*, 1958, **56**, 246–250.

STEVENS, S. S. The direct estimation of sensory magnitudes—loudness. *Amer. J. Psychol.*, 1956, **69**, 1–25.

STEVENS, S. S. On the psychophysical law. *Psychol. Rev.*, 1957, **64**, 153–181.

STEVENS, S. S., CARTON, A. S., & SHICKMAN, G. M. A scale of apparent intensity of electric shock. *J. exp. Psychol.*, 1958, **56**, 328–334.

STEVENS, S. S., & GALANTER, E. H. Ratio scales and category scales for a dozen perceptual continua. *J. exp. Psychol.*, 1957, **54**, 377–411.

(Received March 20, 1958)

Journal of Experimental Psychology
Vol. 57, No. 2, 1959

PAIRED-ASSOCIATES LEARNING AS A FUNCTION OF PERCENTAGE OF OCCURRENCE OF RESPONSE MEMBERS (REINFORCEMENT) [1]

ALBERT E. GOSS, CHURCHILL H. MORGAN, AND SANFORD J. GOLIN

University of Massachusetts

Learning the correct anticipation for each pair of a list of paired associates can be considered analogous to classical conditioning. The stimulus member of a pair corresponds to the CS, the response member and resultant response parallel the UCS-UCR relationship, and arousal by the stimulus member of a response it did not elicit initially is comparable to the formation of a CS-CR association. Because of this correspondence, the two experiments with paired associates reported here—in which the percentage of occurrence of response members (reinforcement) was varied—may be regarded within the framework of studies dealing with the acquisition and extinction of classically-conditioned responses as a function of partial reinforcement schedules (Grant & Schipper, 1952).

Also obtained were data on the effects of post-criterion changes of acquisition percentages of reinforcement to half values or to 0% occurrence. Finally, presentation of members of the pairs at a 2-sec. rate in one experiment and at a 3-sec. rate in the other permitted assessment of the generality of the effects of acquisition and changed schedules of reinforcement. Thus, paired-associates learning and performance were related to acquisition percentage of reinforcement, degree of post-criterion changes of those percentages, and rate of presentation.

Because of the greater number of stimulus-response associations among the stimuli in lists of paired associates, percentage of reinforcement is likely to have different effects on paired-associates learning than on classical conditioning. At present, therefore, the proposed extension from classical conditioning to paired-associates learning should be regarded as mainly procedural in the form of the scheduling of response member homologues of the UCS or UCS-UCR. Although thus limited with respect to interpretation by means of straightforward extrapolation of principles derived from classical conditioning, the data obtained in the present experiments are of immediate interest in that they describe acquisition of paired associates as a function of percentage of reinforcement.

METHOD

Subjects.—The 80 introductory psychology students of one experiment and the 96 of the other were divided into eight groups of 10 and 12 each, respectively. The number of men and women in these groups deviated from a 50–50 split by no more than one. Six additional *Ss* in the first experiment and 13 in the second either had to leave for class before reaching acquisition criteria or were eliminated because of failure to understand instructions, reported inability to concentrate, or *Es'* errors.

Paired associates and apparatus.—The eight paired associates of both experiments were drawn from Underwood's (1952) list

[1] One of the experiments has been reported in greater detail by Morgan (1954). A grant from the Research Council of the University of Massachusetts facilitated completion of the other experiment. Harvey Lifton and Herbert Levitt aided in the statistical analyses and presentation of the data.

of nonsense syllables of intermediate association values and low similarity. They were: *fec-jer, hax-myd, jal-tof, kem-wus, niw-xel, ryg-kon, toq-bih, vod-caz.* Anticipations of response members were spelled out by *Ss*.

Four different random orders of stimulus members of the pairs were used in acquisition and changed percentage phases of both experiments, regardless of the schedules of occurrence of their response members. All lists were typed on tapes in pica capitals and presented on a memory drum which had cam adjustments for 2-sec. or 3-sec. rates.

Scheduling of response members.—Scheduling of response members to achieve 100%, 75%, 50%, and 25% reinforcements involved two restrictions on entirely random scheduling. First, long runs of occurrence or nonoccurrence of response members were avoided by pairing them with stimulus members so that each of the acquisition percentages was realized within each of successive 16-trial blocks. For post-acquisition changes to 37.5% and 12.5% schedules, the blocks were reduced to eight trials. Because four different random orders of stimulus members were used, regardless of schedules of occurrence of response members, more specific limitations on randomness were introduced within this general restriction. These will be noted in connection with subsequent, more detailed description of each percentage schedule.

Second, the percentage of occurrence of response members on each trial was the same as the percentage schedule for each pair; that is, if response members of each pair were on a 75% schedule, six of the eight stimulus members were accompanied by their response members on each trial.

Acquisition.—Paired associates were acquired under 100% occurrence of response members and semi-random schedules of 75%, 50%, and 25% occurrence. For each of these percentages there were two subgroups of 10 *Ss* each at the 2-sec. rate. At the 3-sec. rate, there were two subgroups of 12 *Ss* each for each acquisition percentage.

The 100% schedule was the conventional paired-associates procedure in which stimulus members of each pair were shown alone for 2 or 3 sec. and were always followed by stimulus-response pairs for 2 or 3 sec., respectively. Within the restriction that each one of the four different random orders of stimulus members should appear four times within successive blocks of 16 trials each, the sequence of occurrence of these orders was random.

The 75% schedules for *each* of the four orders of presentation of stimulus members was composed of four different 75% subschedules of response-member occurrence. The first subschedule of the first of the orders of presentation was formed by random elimination of two of the eight response members. These members were then restored and two of the other six members were eliminated randomly to obtain a second subschedule. These two members were restored to the third subschedule from which two of the four previously unchosen response members were eliminated. Elimination of the residual pair of response members produced the fourth subschedule. Thus, in each of these four subschedules for each of the four orders of stimulus presentation, two different response members were eliminated and six of the eight remained; six of the eight response members always occurred on each trial. With four subschedules of each of the four stimulus orders there were 16 order-subschedule combinations. Each of these combinations occurred once within successive different random sequences of 16 trials each. Consequently each response member appeared on 12 of the 16 trials within each block.

For the 50% schedule, four subschedules of response-member occurrence were prepared for each of the four stimulus orders by random elimination of sets of four response members, within the restriction that each response member occurred in two subschedules for each stimulus order and did not occur in the other two. Each of the 16 combinations of stimulus orders and their response subschedules was used once within successive different random sequences of 16 trials each. Thus, each of the eight stimuli was accompanied by its response member on eight, or 50%, of the trials of each 16-trial block. And four of eight, or 50%, of the response members occurred on each trial.

The 25% schedule was obtained by simply reversing the occurrence and nonoccurrence of response members under the 75% schedule. Two of the eight response members occurred on each trial, and each response member appeared on four, or 25% of the trials within successive 16-trial blocks.

The 75%, 50%, and 25% schedules were introduced on the first trials. Accordingly, *Ss* did not have an opportunity to react to all response members until the end of Trial 2 in the case of the 75% and 50% schedules, and until the end of Trial 4 in the case of the 25% schedule. Adjustment for this lack of exposure to all responses on Trial 1 was considered but was rejected as incompatible with the purpose of the study.

TABLE 1

MEANS AND *SD*'S OF TRIALS AND CORRECT RESPONSES TO THE $\frac{7}{8}$ CRITERION
(*N*'s for each reinforcement percentage at the 2-sec. and 3-sec. rates were 20 and 24, respectively)

Acquisition Reinforcement Percentages	Trials				Correct Responses			
	2-sec. Rate		3-sec. Rate		2-sec. Rate		3-sec. Rate	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
100	24.8	8.9	16.4	6.8	59.4	26.5	43.7	19.6
75	34.3	16.8	18.4	5.5	70.2	38.5	44.6	16.8
50	40.6	23.5	20.0	7.0	94.0	59.4	45.6	21.1
25	40.5	12.1	25.3	8.2	102.0	38.8	66.3	24.5

At the 3-sec. rate, paired associates were learned to a criterion of 15/16 correct anticipations on two successive trials. In order to minimize the number of *Ss* failing to finish the experiment within the practical limit of 1 hr., the criterion for the 2-sec. rate was lowered to $\frac{7}{8}$ correct anticipations on a single trial. The intertrial intervals were approximately 12 sec.

The same standard instructions were administered to all groups of both experiments.

Changed percentages.—Upon attainment of the $\frac{15}{16}$ or $\frac{7}{8}$ acquisition criteria, and without further instructions, each of the two subgroups for each acquisition schedule was shifted to 16 trials under a changed reinforcement schedule. The occurrence of response members was completely eliminated for one subgroup and halved for the other subgroup under each acquisition schedule. The arbitrary decision of an earlier study (Goss & Rabaioli, 1952) was the basis for halving the acquisition percentages. Elimination of response members was considered a procedural parallel of the usual condition of extinction of classically conditioned responses.

At each rate the subgroup shifted from 100% to 50% reinforcement was given the first sequence of 16 order-subschedule combinations of the 50% acquisition schedule. Similarly, the first sequence of 16 order-subschedule combinations of the 25% acquisition schedule were used for the 16 trials administered to the subgroup at each rate changed from 50% to 25% reinforcement.

Construction of the 37.5% and 12.5% schedules followed the principles used for the other schedules. There were four orders of occurrence of stimulus members for each of which four subschedules were formed by random elimination of response members. The 37.5% schedule was prepared so that sets of three response members each occurred on each trial, with the restriction that the response member of each pair appeared three times in the first eight trials and three times in the second eight trials.

For the 12.5% schedule, the single response member which was to occur on each of the first eight trials was randomly selected from among those not previously selected. The same selection procedure was used for the last eight trials. Thus each response member appeared once in each half of the 16 trials.

RESULTS

Acquisition.—Acquisition data for the 3-sec. rate were also scored to a $\frac{7}{8}$ criterion and treated along with data for the 2-sec. rate. Means of trials and correct responses to the $\frac{7}{8}$ criterion with 100%, 75%, 50%, and 25% reinforcement schedules at both rates are shown in Table 1 and Fig. 1. Figure 1 also shows the means of reinforced correct responses to the $\frac{7}{8}$ criterion. These are responses which were followed by occurrence of response members. With the 100% schedule, all responses, whether correct or incorrect, were followed by response members. Thus, means for correct and for reinforced correct responses were identical. With the 75% schedule, means of reinforced correct responses can be estimated by multiplying means of correct

responses by .75. With the 50% and 25% schedules, means of reinforced correct responses can be estimated by multiplying means for correct responses by .50 and .25, respectively. Since means of reinforced correct responses can be thus calculated, they were not presented in Table 1.

Within each of the eight combinations of acquisition percentages and rates of presentation there were two subgroups. For one subgroup the acquisition percentage was later halved, while for the other subgroup reinforcement was eliminated. Since assignment of *Ss* to these subgroups was random, no significant differences between the subgroups of each of the percentage-rate combinations was expected for either trials or correct responses to criterion. These subgroups entered the analyses of variance of Table 2 as degree of change. None of the *F*'s involving this variable was significant. For this reason subgroup means for trials and for correct responses were pooled for computations of the means of Table 1.

The other two variables of Table 2 are acquisition percentage of reinforcement and rate of presentation. The significant *F*'s for acquisition percentage indicate that, as the percentages of reinforcement increased from 25 to 100, both trials and correct responses to criterion decreased. Presentation of the stimuli of the paired associates at the 3-sec. rate resulted in significantly lower means of trials and correct responses at each reinforcement percentage. None of the interactions was significant.[2]

[2] Separate analyses of variance for each of the response measures at both 2-sec. and 3-sec. rates yielded *F*'s consistent with those of Table 3. The $\frac{14}{16}$ criterion for the 3-sec. rate resulted in larger over-all values for means for each of the three measures and trends which departed significantly from linearity in the direction of negative acceleration.

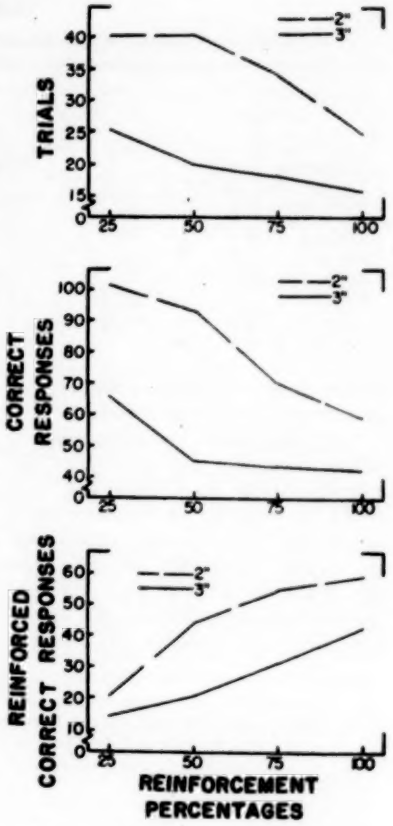

FIG. 1. Means of trials, correct responses, and reinforced correct responses to the $\frac{7}{8}$ criterion at 2-sec. and 3-sec. rates.

Changed percentages.—Table 3 presents means and *SD*'s of correct responses to the $\frac{7}{8}$ criterion trial at the 2-sec. rate and mean correct anticipations per trial for the two $\frac{15}{16}$ criterion trials at the 3-sec. rate. Also shown are means and *SD*'s of correct anticipations per trial within each of the four successive 4-trial

TABLE 2

ANALYSES OF VARIANCE OF TRIALS AND CORRECT RESPONSES TO THE ¼ CRITERION

Source	*df*	Trials		Correct Responses	
		MS	*F*	*MS*	*F*
Acquisition % (AP)	3	1,185.86	8.68*	8,631.86	7.93*
Degree of change (DC)	1	3.27	—	21.12	—
Rate (R)	1	9,829.10	71.99*	42,846.91	39.35*
AP × DC	3	145.65	1.07	1,983.70	1.82
AP × R	3	278.99	2.04	2,132.60	1.96
DC × R	1	5.21	—	157.25	—
AP × DC × R	3	186.81	1.37	1,705.93	1.57
Error	160	136.54		1,088.91	

* $P < .01$.

blocks under changed schedules of occurrence of response members.

The significance of changes in correct responses per trial from the criterion trials to Trials 1–4 was tested by analysis of variance with acquisition reinforcement percentage and degree of change as the other two variables (Table 4). For the 2-sec. rate of presentation, the significant sources of variation were degree of change, criterion trials compared with Trials 1–4, and their interaction. Analysis of the simple effects revealed significant decrements from the criterion trial when the acquisition percentages were halved, but not when response members were eliminated entirely. At the 3-sec. rate, only the *F* for acquisition percentages was significant.

With the exception of the 100–0

TABLE 3

MEANS AND *SD*'S OF CORRECT RESPONSES PER TRIAL FOR CRITERION TRIALS AND FOR BLOCKS OF FOUR TRIALS EACH UNDER CHANGED PERCENTAGES

Rates	Acquisition-Changed Percentages	Criterion Trials		Trials Under Changed Percentages							
				1–4		5–8		9–12		13–16	
		Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
2-sec.	100–50	7.1	0.3	6.2	0.9	6.6	0.6	7.4	0.5	7.2	0.6
	100–0	7.2	0.4	6.7	0.8	6.9	1.1	7.2	1.1	7.4	0.8
	75–37.5	7.4	0.5	6.3	0.7	6.7	0.8	7.3	0.6	7.3	0.7
	75–0	7.2	0.4	7.3	0.4	7.2	0.8	7.3	0.6	7.4	0.8
	50–25	7.1	0.3	6.4	0.8	6.7	0.8	7.2	0.4	7.5	0.5
	50–0	7.2	0.4	7.0	0.4	7.0	0.5	7.1	0.6	7.2	0.7
	25–12.5	7.1	0.3	7.0	0.5	7.3	0.6	7.7	0.3	7.6	0.6
	25–0	7.1	0.3	7.0	1.0	7.2	0.9	7.4	0.7	7.2	0.9
3-sec.	100–50	7.5	0.0	7.1	0.5	7.9	0.2	7.8	0.3	7.9	0.2
	100–0	7.6	0.2	7.7	0.6	7.6	0.7	7.4	1.0	7.2	1.6
	75–37.5	7.5	0.1	7.6	0.5	7.8	0.2	7.7	0.4	7.8	0.2
	75–0	7.6	0.4	7.4	0.3	7.8	0.3	7.8	0.5	7.8	0.6
	50–25	7.8	0.2	7.5	0.7	7.9	0.3	7.9	0.1	8.0	0.0
	50–0	7.5	0.1	7.7	0.4	8.0	0.0	7.9	0.3	7.9	0.2
	25–12.5	7.6	0.2	7.7	0.2	7.9	0.1	8.0	0.0	8.0	0.0
	25–0	7.5	0.1	7.8	0.3	7.8	0.4	7.9	0.3	7.9	0.3

TABLE 4

ANALYSES OF VARIANCE OF CORRECT RESPONSES FOR CRITERION TRIALS AND TRIALS 1–4 UNDER CHANGED REINFORCEMENT SCHEDULES

Source	2-sec. Rate			3-sec. Rate		
	df	*MS*	*F*	*df*	*MS*	*F*
Acquisition percentage (AP)	3	0.66	2.20	3	0.42	3.00*
Degree of change (DC)	1	3.02	10.07**	1	0.13	—
AP × DC	3	0.33	1.10	3	0.36	2.57
Error (*b*)	72	0.30		88	0.14	
Change from criterion (CC)	1	7.65	17.79**	1	0.00	—
AP × CC	3	0.66	1.53	3	0.25	1.39
DC × CC	1	3.03	7.05**	1	0.42	2.33
AP × DC × CC	3	0.44	1.02	3	0.32	1.78
Error (*w*)	72	0.43		88	0.18	

* $P = .05$.
** $P < .01$.

group at the 3-sec. rate, correct responses under both rates of presentation increased from Blocks 1–4 to 13–16. The *F*'s for these increases obtained in analyses of variance of the totals of correct responses in each block were highly significant (Table 5). Blocks of trials at the 2-sec. rate entered into significant interactions with acquisition percentages and degree of change. These interactions were due to greater increases for halved than for eliminated changes from 100%, 75%, and 50% acquisition percentages in contrast with little difference between the 25%–12.5% and 25%–0% groups, neither of which exhibited initial decrements nor marked subsequent increases.

The significant interaction of blocks of trials and degree of change at the 3-sec. rate was due to the upward trend for halved percentage conditions which differed from one of zero slope for elimination of response members. The occurrence of a decrease in correct responses for the 100%–0% and of an increase in correct responses for the 100%–50% group, probably contributed most to the triple interaction.

TABLE 5

ANALYSES OF VARIANCE OF CORRECT RESPONSES UNDER CHANGED PERCENTAGES

Source	2-sec. Rate			3-sec. Rate		
	df	*MS*	*F*	*df*	*MS*	*F*
Acquisition % (AP)	3	27.77	—	3	28.54	2.60
Degree of change (DC)	1	24.20	—	1	4.38	—
AP × DC	3	18.10	—	3	4.25	—
Error (*b*)	72	33.56		88	10.99	
Blocks of trials (BT)	3	112.41	130.71*	3	23.14	13.22*
AP × BT	9	4.22	4.91*	9	1.89	1.08
DC × BT	3	33.23	38.64*	3	7.85	4.48*
AP × DC × BT	9	2.20	2.56*	9	6.61	3.78*
Error (*w*)	216	0.86		264	1.75	

* $P < .01$.

Discussion

Acquisition.—Means of trials and correct responses to criterion at both 2-sec. and 3-sec. rates of presentation decreased as functions of acquisition percentage of response-member occurrence. The relationships for means of reinforced correct responses were inverse.

Goss and Rabaioli (1952) found that a 50% schedule of reinforcement of verbal choices nearly doubled trials and correct choices to criterion. In contrast, the 25% schedule for paired associates only increased those measures by from 50% to 70%. That quartering the 100% schedule would produce only this relatively small reduction in rate of learning had not been anticipated. However, it now seems possible that decreasing the percentages of reinforcement may have introduced differences in a number of other variables or conditions which could have served to increase rate of acquisition of the paired associates. These include: (*a*) amount of response competition per trial, (*b*) amount of opportunity to rehearse per trial, (*c*) time intervals between particular stimulus-response pairs, and (*d*) familiarization with stimulus members.

The $\frac{7}{8}$ or $\frac{15}{16}$ acquisition criteria for 2-sec. and 3-sec. rates, respectively, were reached under all four percentages of reinforcement. These findings are comparable to those for runway and simple choice tasks (Jenkins & Stanley, 1950) in which response strengths also go to approximately equal final levels. But the common final level for the correct anticipations of the paired associates contrasts with the apparent equivalence between asymptotic proportions of response occurrence and percentages of reinforcement which have been obtained with the classically conditioned eyelid response (Grant & Schipper, 1952) and "verbal expectations" (Estes & Straughan, 1954; Grant, Hake, & Hornseth, 1951).

Turning to another aspect of acquisition performance, the more rapid learning under the 3-sec. than under the 2-sec. rate of presentation might be attributed to different *E*'s. In view of Hovland's (1949) findings for 1-sec. and 2-sec. rates, however, it seems more likely that rate differences rather than *E* differences were the basis of the observed superiority of the 3-sec. rate.

Changed percentages.—At the 2-sec. rate, halving acquisition percentages of response-member occurrence or eliminating response members reduced the number of correct responses during the first four trials under the changed percentages to less than the criterion level. These decrements were somewhat greater for halved changes. The failure to obtain such decrements following attainment of the $\frac{15}{16}$ criterion under the 3-sec. rate may have been due to the more stringent criterion, the slower rate of presentation, or to both factors.

Neither acquisition percentages and degree of change, nor their interaction for either rate of presentation, influenced correct responses under the changed percentages. There were significant upward trends over the four blocks of changed percentage trials which, at the 2-sec. rate, were somewhat steeper for larger acquisition percentages of response-member occurrence and for changes to a halved schedule. The latter effect was probably due to the greater initial decrements. The steeper trend for halved changes at the 3-sec. rate may have been due to more rapid further learning under those percentages of response-member occurrence than with elimination of response members.

Since seven of the eight means for the 13–16 block at both 2-sec. and 3-sec. rates were greater than corresponding criterion trial means, *S*s apparently continued to learn under halved or eliminated conditions of response-member occurrence. The conditions mediating such further learning by *S*s for whom response members were eliminated are speculative.

These apparent upward trends for elimination of response member conditions are not consistent with the results of Peak and Deese (1937) who found that 50 trials involving the presentation

of stimulus members alone resulted in incorrect recall or failure of occurrence of about 25% of previously correct responses. Perhaps the apparent upward trends in correct responses under the 0% schedule of this study would have been reversed at some point beyond 16 trials.

Findings different from those described here have also been reported by Peterson (1956) for a selective learning task. She found that elimination of reinforcement by *E* saying "Right" brought about extinction of responses of low and high cultural frequency which had been practiced for 4, 8, or 16 trials under 100%, 50%, or 25% schedules. Rate of extinction was inversely related to number of learning trials and directly related to percentage of reinforcement. Because her task differed from paired-associates learning in many ways, no satisfactory explanation of the differences in results can be advanced.

In other studies, post-acquisition shifts to lower or 0% schedules of reinforcement of classically conditioned and other simple responses have been accompanied by shifts downward to asymptotes approximating the new percentages of reinforcement (Estes & Straughan, 1954; Grant, Hake, & Hornseth, 1951; Grant & Schipper, 1952). No such immediate, monotonic changes occurred for either multiple choices (Goss & Rabaioli, 1952) or paired associates. The decrement which Peak and Deese (1937) obtained for a shift from a 100% schedule to 50 trials of 0% reinforcement was markedly smaller than the reductions of strengths of simple responses brought about by 100% to 0% shifts. It seems probable, therefore, that extinction trends for more complex tasks differ from those for classical conditioning, "verbal expectations," and simple choice tasks.

Summary

Paired-associates performance was investigated as a function of percentage of occurrence of response members of the pairs during acquisition at 2-sec. and 3-sec. rates of presentation and of post-acquisition halving of the percentages of occurrence of response members or their elimination. Eight pairs of nonsense syllables were presented with percentages of occurrence of response members of 100, 75, 50, and 25 during acquisition, and 50 or 0, 37.5 or 0, 25 or 0, and 12.5 or 0 under changed schedules, respectively.

With increasing acquisition percentages, means of trials and correct responses to criterion increased and those of reinforced correct responses to criterion decreased. Learning at the 2-sec. rate was significantly slower than at the 3-sec. rate, but acquisition percentages of reinforcement did not interact with rates of presentation.

At the 2-sec. rate, eliminating response members or halving the percentages with which they occurred produced an immediate decrement in correct responses which was followed by further learning. Probably because of the more stringent acquisition criterion, at the 3-sec. rate no decrements followed such changes. Neither main effects of acquisition percentage of reinforcement and of degree of change in those percentages nor their interaction influenced frequencies of correct responses under changed percentages of reinforcement.

References

Estes, W. K., & Straughan, J. H. Analysis of a verbal conditioning situation in terms of statistical learning theory. *J. exp. Psychol.*, 1954, **47**, 225–234.

Goss, A. E., & Rabaioli, E. J. Response strength in a modified Thorndikian multiple-choice situation as a function of varying proportions of reinforcement. *J. exp. Psychol.*, 1952, **43**, 106–113.

Grant, D. A., Hake, H. W., & Hornseth, J. P. Acquisition and extinction of a verbal conditioned response with differing percentages of reinforcement. *J. exp. Psychol.*, 1951, **42**, 1–5.

Grant, D. A., & Schipper, L. M. The acquisition and extinction of conditioned eyelid responses as a function of the percentage of fixed-ratio random reinforcement. *J. exp. Psychol.*, 1952, **43**, 313–320.

Hovland, C. I. Experimental studies in rote-learning theory: VIII. Distributed practice of paired associates with varying rates of presentation. *J. exp. Psychol.*, 1949, **39**, 714–718.

JENKINS, W. O., & STANLEY, J. C., JR. Partial reinforcement: A review and critique. *Psychol. Bull.*, 1950, **47**, 193–234.

MORGAN, C. H. Paired-associate learning as a function of varying proportions of reinforcement. Unpublished M.S. thesis, Univer. of Massachusetts, 1954.

PEAK, H., & DEESE, L. Experimental extinction of verbal material. *J. exp. Psychol.*, 1937, **20**, 244–261.

PETERSON, M. J. Verbal response strength as a function of cultural frequency, schedule of reinforcement, and number of trials. *J. exp. Psychol.*, 1956, **52**, 371–376.

UNDERWOOD, B. J. Studies of distributed practice: VII. Learning and retention of serial nonsense lists as a function of intralist similarity. *J. exp. Psychol.*, 1952, **44**, 80–87.

(Received March 31, 1958)

Journal of Experimental Psychology
Vol. 57, No. 2, 1959

PREDICTION OF SEQUENTIAL TWO-CHOICE DECISIONS FROM EVENT RUNS [1]

DELMER C. NICKS

San Fernando Valley State College

Probability-learning tasks have been used quite extensively recently as tests of stochastic theories of learning (Bush & Mosteller, 1955; Estes & Straughan, 1954). However, they may also be considered as sequential decision tasks. They are, in effect, decision tasks in which an *S* must learn about the situation by observing the results of his own past decisions (Edwards, 1956). From this point of view individual predictions and the factors related to them are of interest, in addition to the over-all proportions of choices of the alternatives. There is little information available concerning individual choices in two-choice decision tasks. The two most relevant studies are those by Jarvik (1951) and Hake and Hyman (1953). Jarvik found that, during a run of occurrences of one alternative, choices of that alternative dropped rather sharply. He called this the negative recency effect and likened it to the well known "gamblers fallacy." Hake and Hyman constructed sequences in which the probability of occurrence of an alternative when it had occurred on the immediately previous trial was not the same as when the other event had occurred on the previous trial. In other words, the second-order conditional probabilities of occurrence of the events were not the same as the first-order probabilities. Predicting these sequences, *S*s appeared to learn the sequential dependencies; i.e., the frequencies with which they predicted an event varied appropriately with the immediately preceding event. However, in both of these studies the over-all proportions of predictions still matched the proportions of occurrence of the alternatives. It is likely, then, that individual predictions of *S*s are closely related to the structure of the particular sequence being predicted as well as to the over-all proportions of the alternatives.

The investigation of sequence structure as an independent variable immediately presents two problems. First, how much of the sequence must be considered in relation to *S*'s predictions on a given trial and, secondly, what is the best way to describe or define sequence structure. Though there are a number of ways to describe sequence structure, event runs seemed a good place to start in light of the two aforementioned studies. An event run is a series of occurrences of the same event, bounded on either end by the other event or one end of the sequence; i.e., the sequence RR GGG R GG RRR GG contains runs of length 2, 1, and 3 of R events and runs of length 3, 2, and 2 of G events. Describing sequences in terms of the number and

[1] This research was done under ARDC Project No. 7737, Task No. 27107, in accordance with ARDC Technical Requirement No. 181-56, in support of the research and development program of the Air Force Personnel and Training Research Center, Lackland Air Force Base, Texas. Permission is granted for reproduction, translation, publication, use and disposal in whole and in part by or for the United States Government.

lengths of runs of the alternatives has several experimental and conceptual advantages. First, a complete description of the run structure and how it is sampled contains a description of higher-order conditional probabilities as well as of the first-order probabilities. Secondly, within this framework it is possible to describe the over-all sequence structure as well as any given segment preceding a particular response. And thirdly, it seems intuitively reasonable that *Ss* themselves may view a sequence of binary events as a sequence of runs rather than as a sequence of individual events.

The following experiment was designed to investigate the relationship between event runs and *S*'s behavior in a two-alternative decision task.

Method

Apparatus.—A small gray box on *E*'s table contained a red and a green jewel light, side by side, about 2 in. apart. Two switches on the back of the box controlled the two jewel lights. There were six chairs for *Ss* and a small box, open at the front, on the arm of each chair. Each box contained two response buttons. On top of the box there was an "R" over one button and a "G" over the other. The *Ss* recorded their predictions by pushing one of these buttons. The boxes prevented any *S* from seeing what any other *S* was doing. There were two channels for each *S* on an Esterline-Angus 20-channel-operations recorder and two channels for *E*, giving a continuous record of all predictions and events. The stimulus lights were operated by *E* according to a predetermined program.

TABLE 1

Distribution of Run Lengths for All Sequences

Run Length	50:50 Random		50:50 Non-random		67:33		75:25	
	G	R	G	R	G	R	G	R
1	48	48	48	53	29	58	18	55
2	24	24	24	29	18	20	13	14
3	12	12	12	4	13	6	11	3
4	6	6	6	0	10	3	9	1
5	3	3	3	1	6		6	
6	2	2	2	3	4		4	
7	1	1	1	4	2		2	
8				2	2		2	
9					2		2	
10					1		2	
11							1	
12							1	
13							1	
14							1	
Rand. G	190	190	190	190	256	128	285	96
Total	380		380		384		381	

Sequences.—Each sequence was approximately 380 trials long. The proportions of the alternatives in the first three sequences were 50:50, 67:33, and 75:25. Since there was only one sequence of each type, "typical" sequences were constructed in the following fashion. For each sequence, the expected number of runs of each alternative from one on up was computed until the expected number of runs of each alternative was less than .5. This gave the desired run structure. In some cases this distribution was juggled slightly to make the total numbers of runs and of events come out right. However, in no case does the actual number of runs of a given length deviate from the expected number by more than one. These runs were then drawn randomly, without replacement, alternating between red and green runs, until all runs had been drawn, giving the desired sequence. Higher-order conditional probabilities, when checked, turned out to be much closer to first-order probabilities than they did in sequences constructed from events rather than runs of events.

The fourth sequence was the 50:50 nonrandom sequence. There were more very short and very long R runs and fewer intermediate ones than in the 50:50 random sequence. The distribution of G runs was the same as in the 50:50 random sequence. Table 1 shows the run structures and lengths of all sequences.

Subjects.—The *Ss* were 288 Air Force basic trainees and 72 Air Force officers from the pre-flight school at Lackland Air Force Base. They were tested in groups of six.

General procedure.—The *E* informed *Ss*, in general terms, that the purpose of the experiment was to study decision making in a laboratory situation and that the results of the study might affect Air Force classification and selection procedures. The *Ss* were

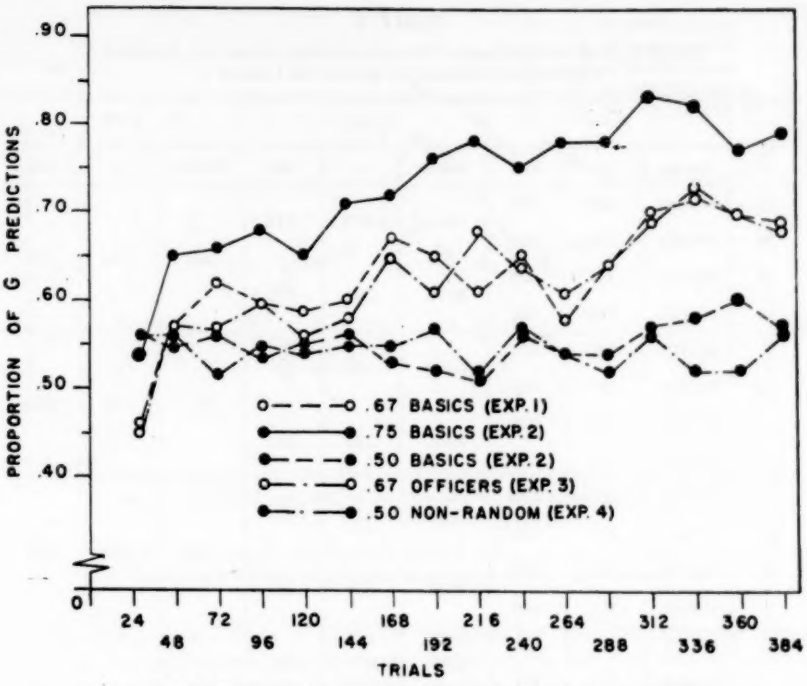

FIG. 1. Probability learning curves in blocks of 24 trials for all groups.

then told that either the red or the green light would go on briefly and that they were to predict which one it would be by pressing one of the two buttons in front of them. After the light went off they were to predict which light would come on next, and so on until the end of the sequence. The *Ss* got one point for predicting correctly which light was to go on and no points if they predicted incorrectly. The objective was to get as many points as possible. However, *Ss* were not allowed to talk or to keep track of their points at any time during the course of the experiment. The *E* told them that there was no pattern and that it would probably hurt their score if they tried to find one. The *Ss* could take as much time as they needed since the next light did not come on until all *Ss* had made their predictions. The *Ss* were told, finally, that there would be quite a few trials and that the experiment would probably take a little over an hour. All groups of *Ss* rather quickly adopted a fairly steady pace of about 12 predictions per minute. On each trial *E* turned on the correct light for about 2 sec. after all *Ss* had made their predictions.

A different group of 72 basic trainees was run on each of the four sequences. The 72 officers were run on the 67:33 sequence to check population differences.

RESULTS

The proportions of green choices in blocks of 24 trials are shown in Fig. 1 for all groups.[2] The curves

[2] The over-all proportion of green predictions for the 50:50 group was .55 rather than .50. A group of 12 *Ss* was run on the mirror image of the sequence (i.e., the alternatives were reversed) and the proportion of green predictions was .56. Another group of 12 *Ss* was run on the original sequence using green

TABLE 2

MEAN PROPORTION OF GREEN PREDICTIONS FOLLOWING ALL POSSIBLE PRECEDING PATTERNS OF EVENTS TO LENGTH 4

N	$n = 4$			$n = 3$			$n = 2$		
	Pattern	p	SE_p	Pattern	p	SE_p	Pattern	p	SE_p
52	GGGG	.636	.016	GGG	.679	.015	GG	.722	.018
30	RGGG	.751	.023						
29	GRGG	.797	.022	RGG	.804	.016			
14	RRGG	.818	.026						
26	GGRG	.736	.023	GRG	.722	.017	RG	.739	.016
16	RGRG	.697	.026						
14	GRRG	.769	.025	RRG	.773	.023			
7	RRRG	.779	.035						
29	GGGR	.451	.024	GGR	.459	.020	GR	.501	.019
13	RGGR	.475	.026						
14	GRGR	.593	.027	RGR	.586	.025			
7	RRGR	.571	.040						
16	GGRR	.511	.028	GRR	.529	.026	RR	.527	.024
5	RGRR	.531	.039						
7	GRRR	.569	.044	RRR	.569	.041			
2	RRRR	.569	.049						

appear to level off somewhat after about 96 trials. In order to study structural effects as independently as possible of probability learning, the analyses that follow are concerned with the last 288 trials unless otherwise specified.

First, how many immediately preceding events are related to prediction on a given trial? There are 2^n possible preceding sequences of events where n is the number of preceding trials considered. Table 2 shows, for the 67:33 basic group, the proportion of G predictions, summed over *S*s and trials, following every possible pattern of preceding events for n equal to 2, 3, and 4. The figure next to RGGG, for example, shows the proportion of green predictions when the fourth previous event was an R light and the other three intervening events were G lights. Hereafter this will be referred to as the RGGG proportion, and so forth. This table also shows the number of times each pattern occurred and the standard errors of the mean proportions.

and white instead of green and red and the proportion of green predictions was .505. It was concluded that the bias was due to a preference for green over red rather than the nature of the sequence itself.

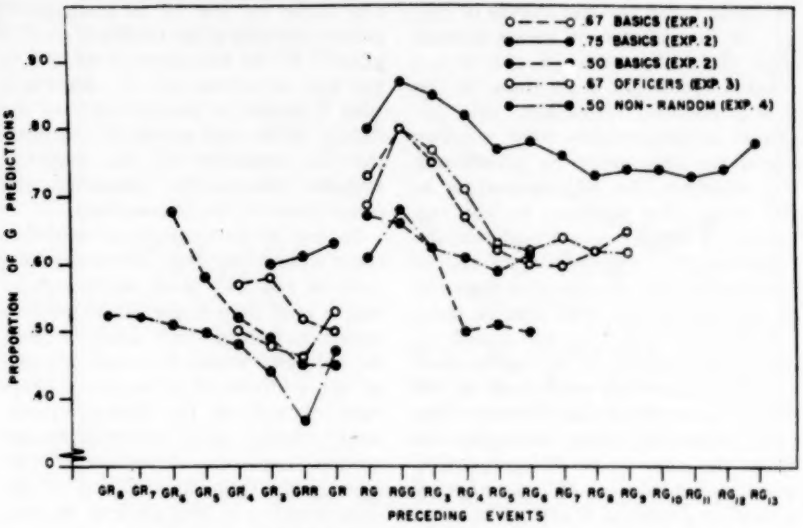

FIG. 2. Run curves for all groups showing the proportion of G predictions following n successive previous occurrences of the same alternative.

The differences in the $n = 4$ column support the hypothesis that the preceding run of like events is related to prediction on a given trial. The difference between the RGGG and GGGG proportions is .11 while the next largest difference between proportions differing only in the fourth previous event is .04. Events this far back in the sequence would appear to be minimally related to prediction except when they occur in a run of like events. Consequently, run curves were constructed for all groups and are shown in Fig. 2. These curves show the proportion of green predictions following 1, 2, 3, etc., immediately preceding green events and 1, 2, 3, etc., immediately preceding red events. The standard errors of these mean proportions are shown in Table 3. (Points are not plotted where a run of that length occurred only once, except for the 75:25 group where many of the long runs were of a unique length. The standard errors of these points are not shown.)

Some interesting comparisons can now be made between the proportions

TABLE 3

STANDARD ERRORS OF MEAN PROPORTIONS IN FIG. 2

Run	50:50 Basics	67:33 Basics	75:25 Basics	50:50 Non-random	67:33 Officers
GR_5 or more	.030			.025	
GR_4	.027	.049		.030	.046
GR_3	.026	.044	.039	.029	.042
GRR	.025	.026	.028	.023	.029
GR	.021	.019	.024	.019	.024
RG	.020	.016	.019	.019	.019
RGG	.022	.016	.015	.021	.017
RG_3	.025	.023	.016	.028	.020
RG_4	.028	.023	.018	.029	.024
RG_5 or more	.031	.022	.022	.030	.027

in Table 2 and the run curves in Fig. 2. At first glance, it seems strange that the differences in the $n = 3$ column are larger than those in the $n = 2$ column. However, this becomes understandable when the run curve for this group is considered. For example, the RG proportion in the table also appears in the run curve. The GG proportion, on the other hand, is a composite of all the other points on the curve to the right of the RG point. The first of these proportions, and the one based on much the largest N, is higher than the RG proportion while most of the other proportions are lower. The GG proportion, then, averages out about the same as the RG proportion since it is made up of a number of dissimilar groups of trials. The same process works in reverse for the comparison of GGG and RGG. The RGG proportion corresponds to the highest point on the run curve and the GGG proportion contains nearly all of the low points. Similarly, the RGGG proportion corresponds to the RG_3 point on the run curve and the GGGG proportion is a composite of all points to the right of this. Run curve analysis, then, seems to give a more reasonable picture of *Ss*' behavior than the type of analysis in Table 2.

Considering Fig. 2, the shape of the run curve appears to differ for sequences with different proportions of the alternatives. For example, green predictions increased substantially going from RG to RGG for the 67:33 group and the 75:25 group whereas they did not increase at all for the 50:50 group. It seemed likely that these differences were related to the different run structures of the various sequences resulting from the different proportions of the alternatives. The run curve for the 50:50 nonrandom group contains some evidence on this point. If the run curve is related to the run structure of the sequence, then it should be possible to vary the nature of the run curve by changing the run structure of the sequence without necessarily changing the proportions of the alternatives.

In the 50:50 nonrandom sequence there were more short and more long runs of red and fewer intermediate length ones than in the 50:50 random sequence. It seemed unlikely that *Ss*' behavior would be much affected by the addition of a few more short runs of red to the comparatively large number which occurred in the random sequence. However, it did seem likely that the addition of the same number of long runs to the few present in the random sequence might affect *Ss*' behavior. Consequently, it was expected that *Ss*' run behavior with respect to red runs would be more like that of the 67:33 or 75:25 group, where there were a number of long runs, than like that of the 50:50 random group. The run curves for both 50:50 groups can be compared in Fig. 2. The proportion of green predictions increases from GR to GRR as it does for the 67:33 and 75:25 groups. The curve also drops much more slowly from the peak and never drops as far as the curve for the 50:50 random group. However, surprisingly enough, the same is also true of the curve for prediction following G runs though the run structure for G events was identical with that in the 50:50 random group. It is possible that this result was due to generalization. If *Ss*' responses are, in some way, related to the run structures for the two types of events, it is not unreasonable to suppose that some generaliza-

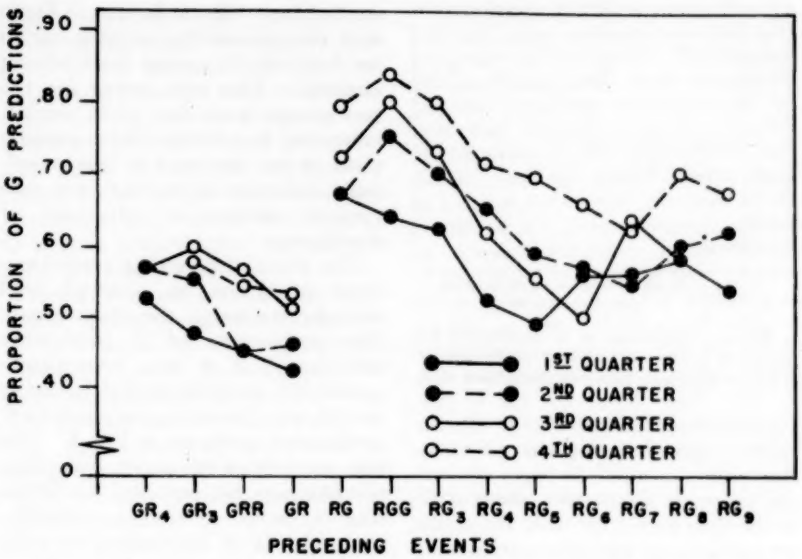

FIG. 3. Run curves by quarters of the sequence for the .67 group of basic trainees.

tion might occur with respect to the two run structures. In sequences with proportions other than 50:50 the expected run structures would not be identical and one might predict progressively less generalization as the proportions deviate from 50:50. However, the present study contains no evidence relevant to this prediction.

In order to investigate the nature of the run curve as a function of trials, run curves were plotted for each quarter of the sequence for the 67:33 basic group. These are shown in Fig. 3. The curve for green runs during the first quarter (96 trials) is somewhat different from the same portion of the curve for the other three quarters. However, the shapes of the curves appear quite similar for the last three quarters, though the level of the curves rises as would be expected from the probability learning curve in Fig. 1. It would appear, then, that behavior with respect to runs changed somewhat during the first quarter for this group. When separate run curves were plotted in this fashion for the 50:50 random group, however, there did not appear to be any change in the shape of the curves. Also, the run curve for the first quarter of the 67:33 basic group was quite similar to the run curve for the 50:50 group, the higher level of the 67:33 curve reflecting the probability learning which took place in the first quarter. During the initial part of the sequence, then, behavior with respect to event runs was similar for the 67:33 and 50:50 groups, but the run curves changed for the 67:33 group during the latter part of the sequence whereas they did not for the 50:50 group.

When *Ss* predict binary sequences they behave as though they had

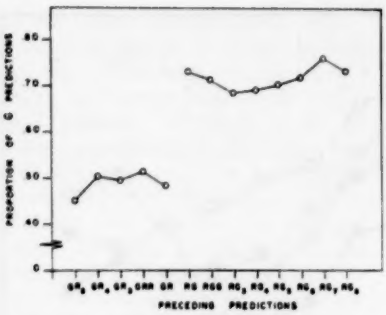

FIG. 4. Proportion of G predictions following n successive previous predictions of G and n successive previous predictions of R.

certain expectations about the sequence upon entering the situation. For example, most *Ss* begin by predicting each alternative about 50% of the time no matter what the actual proportions of the alternatives may be. If the proportions are other than 50:50, *Ss* predict the more frequent alternative oftener as the sequence continues. If the sequence is a 50:50 one, their behavior does not change in this respect. Predictions in connection with event runs may behave in a similar manner. It would appear that *Ss* also bring to the situation certain a priori expectations concerning the structure of the sequence in terms of runs. Since run curves did not change over trials for the 50:50 random sequence, one might say these expectations were confirmed in some manner by the run structure of the sequence. This is not to say, of course, that *Ss'* expectations were correct. As a matter of fact, it is quite possible that the inaccuracies in these expectations are what give the run curve its characteristic shape.

The 67:33 officer group was included as a check upon population differences in the type of behavior studied here. It can be seen in Fig. 1 that the probability learning curves for both 67:33 groups were almost identical. The run curves for the two groups were also quite similar, indicating that in two-choice decision tasks of the sort used in this experiment, behavior is probably not particularly related to education or intelligence.

The effect of preceding predictions upon prediction on a given trial was also studied in the officer group. The proportions of G predictions following 1, 2, 3, etc., immediately previous G predictions and following 1, 2, 3, etc., immediately previous R predictions are shown in Fig. 4. The two portions of the curve, depending on whether the previous prediction was R or G, differed considerably in level. It is interesting to note,

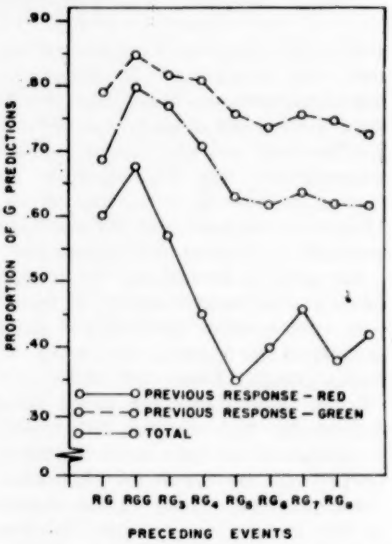

FIG. 5. Run curves for the .67 officer group, plotted separately for those *Ss* predicting G on the previous trial and for those *Ss* predicting R on the previous trial.

incidentally, that the proportions following previous R predictions were much farther from the over-all level of G prediction than were the proportions following previous G predictions. However, each portion of the curve was fairly horizontal, particularly during the initial portion. The effect of previous predictions, then, appears to depend mainly upon the immediately previous prediction.

In Fig. 5 the run curve for the 67:33 officer group is plotted separately at each point for those predictions preceded by a G prediction and for those preceded by an R prediction. The over-all run curve is also shown. The two curves differ in level as would be expected from Fig. 4. The shape of the two curves is similar at the beginning but the run curve for G predictions following an R prediction drops rapidly whereas the run curve following G predictions drops very slightly. One could tentatively infer that the previous prediction interacts with the run curve only slightly at the beginning of a run but progressively more as the length of the run increases. This could be caused by several factors or might be an artifact of these particular curves. The interpretation of this result is tenuous since there is a possible *S* artifact. It is likely that "green responders" may be overrepresented in one curve while "red responders" are overrepresented in the other.

In terms of binary decision making, the preceding run of like events appears to be definitely related to prediction on a given trial. It is possible that other preceding patterns of events are also related to prediction, but the proportions in the $n = 4$ column of Table 2 suggest that these other effects are fairly minor.

Concerning decision-making theories, one main point emerges from the data presented in this study. The relation between a particular event occurrence and the probability of prediction on the next trial depends upon the position of the event in the sequence—in particular, upon the position of the event in a run of like events. In other words, prediction of a given event may increase or decrease following the occurrence of that event, depending upon whether the occurrence was at the beginning of a run or not. Decision theories as usually applied to this type of task do not predict this result, though this is not to say that they could not. It should be worthwhile to explore the possibility of manipulating the variables in some of these theories in such a way as to predict this result.

Summary

This study was an investigation of trial-by-trial behavior in sequential two-choice decision tasks, or probability-learning tasks, as a function of event runs. A different group of 72 *Ss* was used with each of four different sequences. The proportions of the alternatives in three random sequences was 50:50, 67:33, and 75:25. The fourth sequence was a 50:50 nonrandom sequence. Run curves were plotted for each group showing the proportion of predictions of one event after n successive previous occurrences of that event and after n successive previous occurrences of the other event. The typical run curve rose following the first two occurrences of the event being predicted and then dropped gradually, leveling off after six or seven successive occurrences of that event.

The shape of the run curve was, partially at least, dependent upon the run structure of the sequence. Consequently, the shape of the curve changed when the over-all proportions of the alternatives varied and also when the structure of the event sequence was modified independently of the proportions of the alternatives. Apparently, run curves reflect an important aspect of *Ss'* behavior which does not appear in probability-learning curves.

References

BUSH, R. R., & MOSTELLER, F. *Stochastic models for learning*. New York: Wiley, 1955.

EDWARDS. W. Reward probability, amount, and information as determiners of sequential two-alternative decisions. *J. exp. Psychol.*, 1956, **52**, 177–188.

ESTES, W. K. & STRAUGHAN, J. H. Analysis of a verbal conditioning situation in terms of statistical learning theory. *J. exp. Psychol.*, 1954, **47**, 225–234.

HAKE, H. W., & HYMAN, R. Perception of the statistical structure of a random series of binary symbols. *J. exp. Psychol.*, 1953, **45**, 64–74.

JARVIK, M. E. Probability learning and a negative recency effect in the serial anticipation of alternative symbols. *J. exp. Psychol.*, 1951, **41**, 291–297.

(Received March 20, 1958)

Journal of Experimental Psychology
Vol. 57, No. 2, 1959

PERFORMANCE AS A JOINT FUNCTION OF AMOUNT OF REINFORCEMENT AND INTER-REINFORCEMENT INTERVAL [1]

GEORGE COLLIER AND MAURICE SISKEL, JR.

University of Missouri *University of Nebraska*

One of the keys to the nature of reinforcement is the relationship between performance and amount of reinforcement. The present study is concerned with the character of this relation as a function of inter-reinforcement time when reinforcements are presented on a periodic schedule.

A number of different primary events have been associated with the occurrence of reinforcement. Drive reduction, the consummatory response, receptor stimulation, affective processes, and cognition have all been proposed. Each of these hypotheses has been thought to require an increasing function relating amount of reinforcement and response strength, and this requirement has been supported by a number of experiments (e.g., Crespi, 1944; Guttman, 1953, 1954; Hutt, 1954; Young & Shuford, 1954). No one of these hypotheses has embodied an account of the temporal properties of reinforcement.

Under some circumstances the amount of reinforcement function is nonmonotonic, i.e., showing a decrease in response strength for larger amounts of reinforcement. Explanations involving adaptation, satiation, aversion, and consummatory activity have been proposed to account for this nonmonotonicity (e.g., Guttman, 1953; Hellyer, 1953; Hughes, 1957; Pfaffman, 1952). It would be logical to expect on the basis of these explanations that the extent of nonmonotonicity would be some function of the time interval between successive reinforcements. It is the aim of this study to examine this possibility and to study the intrinsically interesting character of the interaction between reinforcement and inter-reinforcement interval.

METHOD

Apparatus.—Four Skinner boxes were used which have been described in detail by Guttman (1954). They delivered .02 ml., ±.002 ml., of fluid reward. Temperature and relative humidity were controlled by air conditioning equipment. Number of bar-press responses was recorded cumulatively at 5-min. intervals.

Subjects.—The *S*s were 16 male albino rats, 77–80 days old, obtained from the Duke University Medical School colony.

Procedure.—The *S*s were placed on a 22-hr. food and water deprivation schedule. Two days of magazine training were followed by acquisition of the bar-pressing response under continuous water reinforcement until all *S*s had given approximately 800 responses. The 64 experimental days followed. Throughout experimentation, *S*s were maintained on a 23-hr. food privation schedule and ad libitum water.

Experimental design.—Four amounts of reinforcement (4%, 8%, 16%, and 32% sucrose solutions) and four periodic reinforcement (PR) intervals (.5, 1, 2, and 4 min.) were combined factorially. One *S* was placed in each of the 16 treatment combinations (cells). After four days, 45 min. per day, in its assigned cell, each *S* was placed in a new treatment combination. This procedure was repeated 16 times so that all *S*s performed in all 16 possible combinations. The order of

[1] This investigation was performed at Duke University. It was supported in part by research grants M-631, Duke University, and M-817, University of Missouri, from the National Institute of Mental Health, of the National Institutes of Health, Bethesda, Maryland. The authors wish to acknowledge the contribution of Norman Guttman from whose work this investigation stems.

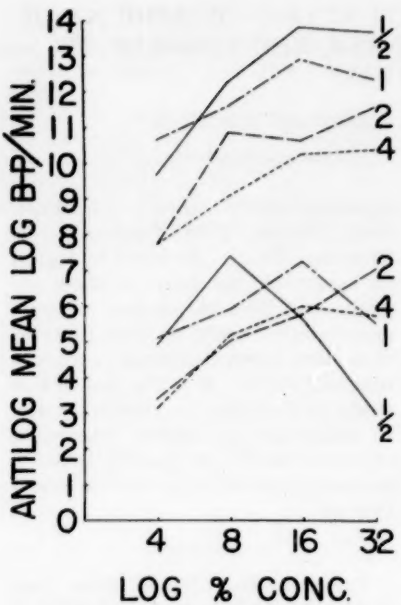

FIG. 1. Rate of bar pressing as a function of concentration and PR interval for the first and last 5 min. of session.

treatments was determined by 16 orthogonal Latin squares.

All sucrose solutions were prepared as percentage by weight. Commercial granulated sugar and distilled water were used.

RESULTS

Figure 1 presents the antilogs of the mean log rates of responding during the first and last 5 min. of the last (fourth) session in each treatment combination. Each curve represents rate of responding plotted against concentration for one of the PR intervals. An analysis of variance on log rate of responding[2] for concentration (C), interval (I), portion of the session (P), and *Ss* (S) is presented in Table 1. Since the triple interactions involving *Ss* did not appear to differ from one another, they were pooled with the quadruple interaction and used as the error term. The main effects of C, I, and P, as well as the C × I and C × I × P interactions, were significant. A similar analysis of the log *total* number of responses per session again found main effects of C and I and the C × I interaction significant.

As is to be expected, the rate of responding is higher for the greater concentrations and shorter intervals, and lower at the end of a session than at the beginning. The interactions, on the other hand, of C, I, and P are very interesting. Examination of Fig. 1 indicates that the nonmonotonicity of the amount of reinforcement function is a decreasing function of the inter-reinforcement interval and an increasing function of the duration of the session.

In order to examine the interaction effects in more detail free of main effects, twice corrected tables (Lind-

[2] As is usual with bar-pressing data at this level of responding, it was necessary to use a log transformation to obtain homogeneity of variance.

TABLE 1

ANALYSIS OF VARIANCE OF LOG RATE OF BAR PRESSING

Source	*df*	*MS*	*F*
C	3	.59496	11.22*
I	3	.23265	4.39*
P	1	13.04487	99.35*
S	15	.87981	16.59**
C × I	9	.12367	2.33**
C × P	3	.06598	1.24
C × S	45	.03787	.71
I × P	3	.07993	1.51
I × S	45	.04555	.86
P × S	15	.13130	2.48*
C × I × P	9	.10169	1.92**
C × I × S	135	.04459	.57
C × P × S	45	.03012	.38
I × P × S	45	.02518	.32
C × I × P × S	135	.07838	
Pooled error	360	.05303	

* $P < .01$.
** $P < .05$.

quist, 1953) were constructed for the rate vs. log concentration and rate vs. interval functions. For clarity of presentation, the .5- and 1-min. and the 2- and 4-min. curves for the rate vs. log concentration function and the 4% and 8% and the 16% and 32% curves for the rate vs. interval function were averaged (Fig. 2). From the left-hand side of Fig. 2 it can be seen that the slope of the rate vs. log concentration function is a decreasing function of concentration for short intervals and an increasing function for long intervals. The rate vs. concentration function itself becomes significantly nonmonotonic for the shorter intervals. Similarly, the *slope* of the rate vs. PR interval function (Fig. 2, right-hand side) is a decreasing function of interval up to 2 min. for low concentrations and an increasing function of interval up to 2 min. for high concentrations. The increment or decrement in response rate attributable to an increase in amount of reinforcement is a function of the interval between reinforcements; and conversely, the decrement in response rate attributable to an increase in the interval between reinforcements is a function of the amount of reinforcement being given. The magnitude of these interactions is an increasing function of the duration of the session (compare the early and late portions of Fig. 1).

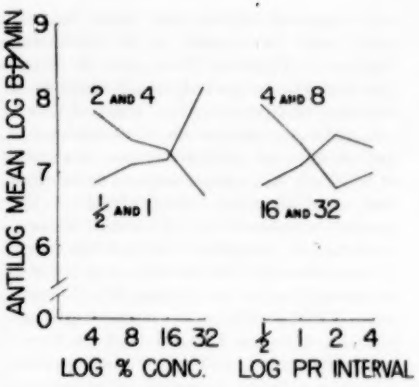

FIG. 2. The interaction between concentration and PR interval.

DISCUSSION

The present study shows that rate of responding is an increasing function of concentration and a decreasing function of the length of the experimental session. This latter variable is, without a doubt, equivalent to the number of successive reinforcements. There is a time parameter in these two functions, the inter-reinforcement interval, which has at least two major effects. First, the *immediate* effects of concentration on rate of responding are modified by the inter-reinforcement interval; low concentrations are more effective than expected if delivered close together, while high concentrations are less effective if delivered close together. Second, the long term, i.e., post-ingestive, effects of concentration are a function of inter-reinforcement interval; the rate at which *S*'s responding slows down is greater for high concentrations at short inter-reinforcement intervals, even though *S* is still consuming all of his reinforcements.

These properties of the amount of reinforcement function may be accounted for by consideration of two classes of events, the stimulation at the time of ingestion, and the average concentration of substance in the alimentary tract. Vigor or frequency of performance, however it is measured, is an increasing function of the amount and intensity of stimulation and a decreasing function of post-ingestive concentration. Concentration of the reinforcing substance corresponds to the intensity dimension of the stimulational process and volume of reinforcing substance to the areal dimension. On these assumptions it would be expected that the initial rate of responding would be an increasing function of concentration, or volume, up to the point at which the effect of the post-ingestive concentration balanced the

stimulational effect, and that beyond this point it would be a decreasing function. Further, the point at which the inversion occurred would itself be a function of concentration, volume, interval between successive reinforcements, and number of reinforcements (duration of session), since these factors determine the post-ingestive concentration. By proper combinations of these values, it should be possible to obtain the point of inversion at different values of volume or concentration (see Crespi, 1944; Guttman, 1953, 1954) or even to obtain no relation between amount and performance (Reynolds, 1952).[3] It is assumed that the relation between vigor or frequency of response and stimulation is itself modified by the quality of the substance consumed and the deprivation state prevailing at the time, while the relation between post-ingestive concentration and response vigor or frequency is a function only of the state of deprivation.

A variety of measures have been found to be increasing functions of concentration of the reinforcing substance. The volume of singly presented sugar solutions consumed in a given interval of time increases up to approximately 5% concentration and then declines (McCleary, 1953; Richter & Campbell, 1940; Young, 1949). A similar relation is found for saccharine (Hughes, 1957). On the other hand, the *initial* rate of ingestion of fluids is a monotonic, increasing function of concentration (Shuford, 1957)[4]—a finding similar to the results of the comparison of the early and late portion of sessions in the present study. A number of studies (e.g., Young & Greene, 1953) show that relative volume consumed of simultaneously offered solutions is an increasing function of concentration, and amount of solute consumed is an increasing function of the highest concentration offered. Frequency of choice is an increasing function of concentration (Young & Greene, 1953). Rate of bar pressing (Guttman, 1953, 1954; Hutt, 1954) and runway velocity (Spence, 1956) increases with concentration. These are linear functions of log concentration up to some maximum value. The relation between log volume and performance is also linear (Hutt, 1954), though this function may be complicated by ingestive properties when solid substances are used. Gustatory threshold studies show concentration is homologous with the intensity of stimulation (Pfaffman, 1952). The higher thresholds and increased reaction times (Holway & Hurvich, 1938) found for stimulation of small areas suggest the homology between the areal properties of stimulation in other receptors and the volume of reinforcement on the gustatory receptors.

The relation between performance and other stimulus parameters of reinforcement is shown in a number of different kinds of studies. Manipulation of either quality or quantity of reward in preference studies results in changes in performance as measured by rate of bar pressing, amount consumed, or frequency of choice (Hutt, 1954; Young, 1949; Young & Shuford, 1954). The short latency of such changes precludes accounting for them by any post-ingestive process. The existence of absolute (Guttman, 1953, 1954; Hutt, 1954; Young, 1949) and differential thresholds, of equal sweetness contours, as well as the classic logarithmic relation between both concentration and volume and performance, all suggest typical psychophysical relations. Variation of performance as a function of variations in amount, concentration, and kind of

[3] These relations are obviously modified by experience as can be simply shown in preference or secondary reward studies, or in a comparison of amount of reward curves obtained using each *S* as its own control with curves obtained using independent groups.

[4] The apparent contradiction between these data and Young's is easily explained since for a high concentration the rate of ingestion, while initially very high, falls off rapidly. Young's points are taken at 10-min. intervals, and by the end of 10 min. at a high concentration *S* has already consumed the amount of sugar he will "accept" at the concentration, an amount above that for lower concentrations, and virtually ceased drinking.

non-nutritive substance (Hughes, 1957; Miller, 1957; Sheffield, Roby, & Campbell, 1954) is further evidence for the control of performance by the reinforcing stimulus.

On the other hand, the existence of satiation, and the data of the stomach loading studies, demonstrate that there is a post-ingestive factor governing the relation between reinforcement and performance (McCleary, 1953; Miller, 1957; Smith & Duffy, 1955, 1957a, 1957b). The time parameters of this process are indicated in the present study (where the rate of bar pressing is a joint function of the amount of substance consumed, the time it takes to get it, and its concentration) and the sugar studies of McCleary (1953) and Shuford (1957). Replotting of McCleary's volume data in terms of amount of sugar consumed shows that amount of sugar consumed is a linear, increasing function of the concentration and that the time in which this amount is consumed is a decreasing function of the concentration. A similar relation for saccharine can be determined from data presented by Hughes (1957) or Carper (1953), a nonmonotonic function for volume drunk, but an increasing monotonic function for amount of saccharine consumed. These results again suggest a post-ingestive mechanism based on momentary post-ingestive concentration. The apparent "overshooting" observed in the amount of sugar consumed may be the result of the emptying of the alimentary tract or of inertia of the mechanism. Further, the rate of ingestion is jointly a function of concentration and time since drinking begins (Shuford, 1957). This interaction between amount of substance ingested, the rate of its ingestion, and its concentration shows that there is a critical time period, governed by concentration, before the amount consumed modifies the relation between performance and reinforcement. The present results support this interpretation. The data of Smith and Duffy (1955), who determined the effects of stomach loading of various sorts of substances on bar-press rate, indicate that the effect is specific to the rewarding substance. They also show that for sugars the shutoff effect appears to be independent of the kind of sugar and only dependent upon the concentration. On the other hand, saccharine appears to have a more immediate shutoff for a much lower concentration (Carper, 1953; Hughes, 1957). This anomaly may be accounted for in terms of some ingestive effect based on adaptation. There is a strong suggestion that the major dimension of the post-ingestive events is osmotic pressure. The means by which this or some other such consequence is reflected back upon consummatory responses and their associated instrumental responses is not obvious. It seems possible that other associated consequences, such as distension, change in body temperature, breathing rate, etc., can become conditioned stimuli for the "shutoff" mechanism (Smith & Duffy, 1957b).

The striking parallelism between functions established by psychophysical and physiological studies of the gustatory process and those found to describe the dependence of performance upon volume and concentration of reinforcement in behavioral studies leads to the conclusion that the major parameters of the amount of reinforcement function, however measured, are intensity, area, and quality of stimulation. Furthermore, an account of satiation and associated phenomena requires a second mechanism, slower acting in nature, which modifies the relation between performance and reinforcement as a function of the momentary post-ingestive concentration of reinforcing substance.

Summary

Rate of bar pressing is examined as a joint function of sucrose concentration, inter-reinforcement (PR) interval, and portion of session.

The interactions as well as the main effects of these variables prove significant. In the first portion of a session performance is an increasing function of concentration for long PR intervals and a nonmonotonic function for short intervals. The point of in-

version occurs at a lower concentration for each interval in the last portion of a session.

It is argued from these data that amount of reinforcement is an increasing function of stimulation and a decreasing function of the post-ingestive concentration of the reinforcing substance.

References

CARPER, J. W. A comparison of the reinforcing value of a nutritive and a non-nutritive substance under conditions of specific and general hunger. *Amer. J. Psychol.*, 1953, **66**, 270–277.

CRESPI, L. P. Amount of reinforcement and level of performance. *Psychol. Rev.*, 1944, **51**, 341–357.

GUTTMAN, N. Operant conditioning, extinction, and periodic reinforcement in relation to concentration of sucrose used as a reinforcing agent. *J. exp Psychol.*, 1953, **46**, 213–224.

GUTTMAN, N. Equal-reinforcement values for sucrose and glucose solutions compared with equal sweetness values. *J. exp. Psychol.*, 1954, **47**, 358–361.

HELLYER, S. Duration of the consummatory response as a variable in amount of reinforcement studies. Unpublished Ph.D. thesis, Indiana Univer., 1953.

HOLWAY, A. H., & HURVICH, L. M. On the psychophysics of taste. I. Pressure and area as variants. *J. exp. Psychol.*, 1938, **23**, 191–198.

HUGHES, L. H. Saccharine reinforcement in a T-maze. *J. comp. physiol. Psychol.*, 1957, **50**, 431–435.

HUTT, P. J. Rate of bar pressing as a function of quality and quantity of food reward. *J. comp. physiol. Psychol.*, 1954, **47**, 235–239.

LINDQUIST, E. F. *Design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin, 1953.

MCCLEARY, R. A. Taste and post-ingestive factors in specific-hunger behavior. *J. comp. physiol. Psychol.*, 1953, **46**, 411–421.

MILLER, N. E. Experiments on motivation. *Science*, 1957, **126**, 1271–1278.

PFAFFMAN, C. Taste preference and aversion following lingual denervation. *J. comp. physiol. Psychol.*, 1952, **45**, 393–400.

REYNOLDS, B. Resistance to extinction as a function of drive reward interaction. *J. comp. physiol. Psychol.*, 1952, **45**, 36–42.

RICHTER, C. P., & CAMPBELL, K. H. Taste thresholds and taste preferences of rats for five common sugars. *J. Nutrit.*, 1940, **20**, 31–46.

SHEFFIELD, F. D., ROBY, T. B., & CAMPBELL, B. A. Drive reduction versus consummatory behavior as determinants of reinforcement. *J. comp. physiol. Psychol.*, 1954, **47**, 349–354.

SHUFORD, E. H., JR. Acceptability and osmotic pressure as determinants of the intake of glucose and sucrose solutions. Paper read at Midwest. Psychol. Ass., Chicago, 1957.

SMITH, M., & DUFFY, M. The effects on intragastric injection of various substances on subsequent bar-pressing. *J. comp. physiol. Psychol.*, 1955, **48**, 387–391.

SMITH, M., & DUFFY, M. Evidence for the dual reinforcing effect of sugar. *J. comp. physiol. Psychol.*, 1957, **50**, 242–247. (a)

SMITH, M., & DUFFY, M. Some physiological factors that regulate eating behavior. *J. comp. physiol. Psychol.*, 1957, **50**, 601–608. (b)

SPENCE, K. W. *Behavior theory and conditioning.* New Haven: Yale Univer. Press, 1956.

YOUNG, P. T. Food-seeking drive, affective processes, and learning. *Psychol. Rev.*, 1949, **56**, 98–121.

YOUNG, P. T., & GREENE, J. T. Quantity of food ingested as a measure of relative acceptability. *J. comp. physiol. Psychol.*, 1953, **46**, 288–294.

YOUNG, P. T., & SHUFORD, E. H., JR. Intensity, duration, and repetition of hedonic processes , related to acquisition of motives. *J. comp. physiol. Psychol.*, 1954, **47**, 298–305.

(Received April 11, 1958)

Journal of Experimental Psychology
Vol. 57, No. 2, 1959

SPONTANEOUS REGRESSION AND RECOVERY IN A SEQUENCE OF DISCRIMINATION PERIODS[1]

DAVID McCONNELL[2]

Indiana University

Estes (1955) has proposed an explicit mathematical model for spontaneous recovery and regression. The present study was designed to test directly certain theoretical consequences which followed from that paper. The model may be briefly recapitulated as follows:

> The total population $S_\bullet$ of stimulus elements can be conceptually partitioned into two mutually exclusive and exhaustive subsets during any experimental period. The set S, called the *available* set, contains those elements attendant during the period, and hence amenable to conditioning or extinction during the period. The *unavailable* set S′ contains all those elements not in S during the training period. On S and S′, respectively, are defined the probabilities $p(t)$ and $p'(t)$ representing the proportions of conditioned elements in the two subsets at time t. During rest intervals between periods a random exchange of elements takes place, whose rate is expressed by the probabilities j and j'. These two parameters are assumed roughly constant in rest intervals, whence the relative size of S with respect to $S_\bullet$ during rest intervals is related to j and j' by the expression $J = \frac{j'}{j + j'}$. From a more rigorous formulation of these assumptions, Estes derived his basic finite difference equation for spontaneous recovery and regression during rest intervals. In the derivation he made the simplifying assumption that no exchange takes place within experimental periods.

From the basic equation the present E derived mathematical predictions for response probability as a function of time following each of a proposed sequence of experimental periods.[3] Figure 1 shows these predictions graphically for four theoretical groups, each of which is subjected to an identical sequence of experimental periods, but no two of which have the same combination of the two rests intervening between experimental periods. The experimental sequence depicted consists of (*a*) acquisition training from $p = 0$ to 1 (where p is proportion of conditioned elements in S), (*b*) extinction to $p = 0$, and (*c*) reconditioning to $p = 1$. The indicated rest intervals are either 0 or ∞. The mathematical predictions on which the graphs are based may be verbally summarized as follows:

1. Prolonging the interval t_1 between Periods I and II should decrease $p(t_1)$ as in any conventional retention study. That is, $p(t_1)$ for Groups 3 and 4 should be less than that for Groups 1 and 2. This is the case because the longer the interval following an initial learning period the more conditioned elements will have escaped into the unavailable set S′.

2. When the interval t_1 between Periods I and II is zero, increasing the interval t_2 between Periods II and III should have no effect on $p(t_2)$ or $p(t_3)$. That is, $p(t_2)$ and $p(t_3)$ for Groups 1 and 2 should be equal. This is the case because all elements conditioned in Period I should be

[1] This study is based upon a part of a dissertation submitted in partial fulfillment of the requirements for the Ph.D. degree in the Department of Psychology, Indiana University, 1957.

[2] Now at Ohio State University.

[3] Readers interested in the mathematical details as well as details of procedure and results may obtain microfilms of the original thesis from University Microfilms, Ann Arbor, Mich.

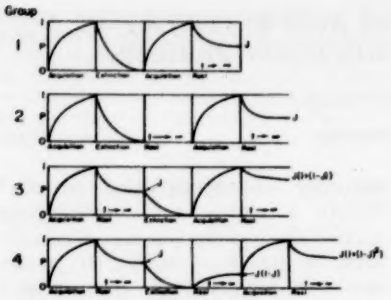

FIG. 1. Predictions of response probability in the available set S for four treatment groups. J is a proportionality constant relating the average size of S to the size of the population, S_*.

extinguished in Period II if no time is permitted after acquisition for escape of elements into S′. Therefore, increasing the recovery time t_2 following extinction in Period II will bring back no additional elements into S.

3. When the interval t_1 between Periods I and II is greater than zero, increasing the interval t_2 between Periods II and III should increase $p(t_2)$ and decrease $p(t_3)$. That is, $p(t_2)$ for Group 4 should be greater than it is for Group 3, but $p(t_3)$ should be less for Group 4 than for Group 3. This is the case because, since the interval between I and II is greater than zero, some elements will escape into S′ before extinction commences in Period II. At the end of II, increasing t_2 will permit more of these conditioned elements to return to S, thus increasing $p(t_2)$ in Group 4. On the other hand, the more of them that are in S at t_2 when Period III reconditioning begins, the fewer will remain in S′ during the period. Since probability in S presumably proceeds to 1, regardless of number of conditioned elements present at the outset, this can only result in a lower over-all number of conditioned elements in the total set S*. Hence, $p(t_3)$ should be lower for Group 4.

4. When the interval t_2 between Periods II and III is zero, increasing the interval t_1 between Periods I and II should increase $p(t_3)$. That is, increasing t_1 increases the number of elements which escape into S′ between Periods I and II. This number will remain fixed until the completion of reconditioning in Period III, if the interval t_2 between II and III is 0. Thus, $p(t_3)$ will be maximized for the longest interval t_1, and we should expect $p(t_3)$ to be higher for Group 3 than for Group 1. Since Groups 1 and 2 should be equal, the comparison should also hold for Groups 3 and 2.

5. When the interval t_2 between Periods II and III is not zero, increasing the length of the previous rest interval t_1 should increase $p(t_2)$ and $p(t_3)$. This means that both $p(t_2)$ and $p(t_3)$ should be higher for Group 4 than for Group 2. Again the relation should hold also between Groups 4 and 1, because of the presumed equality of Groups 1 and 2. As indicated in Prediction 1 above, increasing the first rest interval t_1 should increase the number of conditioned elements escaping from S between Periods I and II. After extinction in Period II, this should leave an increased number to return to S during the recovery interval t_2, and again during the regression interval t_3 following Period III.

Since under the foregoing assumptions there is no intraperiod exchange of elements between S and S′, it is implied that initial ordering in response probability among groups subjected to varying interperiod rest intervals should hold across an entire period following a rest; whence it should not be difficult to check the predictions if differences can be maxi-

mized by employing sufficiently long and sufficiently short rest intervals between periods for different groups. It is evident, in the case of the interval t_3 following Period III, that none of the predictions are testable at $t_3 = 0$, since at that value all probabilities in S should be 1, regardless of previous treatment. Therefore, in the following, t_3 is maximized for all treatment groups, while t_1 and t_2 serve as the two independent variables.

METHOD

Subjects.—The *S*s were 96 female albino rats, age 60–90 days, housed four to a living cage during the experiment.

Apparatus.—Four two-bar Skinner boxes were used. At one end of the box, and 4 in. apart, two ¼-in. brass bars were situated symmetrically with respect to the midline; a .0125 ml. dipper was centered on the midline at the opposite end. Each bar projected ⅜ in., at a height of 1¼ in., from a vertical slot in the wall, and was retractable through the slot. When *S* pressed the reinforced bar, both bars retracted and the dipper descended into a water trough which lay beneath a Plexiglas shield. When *S* made its way to the dipper and stepped on a floor panel in front of the dipper, the bars dropped back into position and the loaded dipper came up to *S*. A 1-in. barrier was erected on the floor in front of the dipper area, two-thirds of the way from the bars, to keep *S*'s tail off the floor panel during bar pressing. When *S* pressed the unreinforced bar, only the recorder was activated. Recording of both reinforced and unreinforced bar presses was accomplished by means of an Esterline-Angus operations recorder.

Pretraining.—All *S*s were given extensive pretraining to one side. Although equal numbers of *S*s in each treatment group were pretrained on each side, we will adopt here the convention, for simplicity, of referring to the pretraining and subsequent even-numbered periods as "R" periods, and to odd periods as "L" periods. About every third *S* entered in pretraining was discarded for failure to meet a fixed rate-of-response criterion during pretraining.

Experimental procedure.—Periods I and III were acquisition periods (Table 1). During these periods, the dipper and retraction circuits closed only for the left (L) bar. During Periods II and IV (extinction) only the R bar operated these circuits. A continuous record was kept of both reinforced and unreinforced responses. The number of trials (reinforced responses) in each period was 30, 40, 40, and 50, respectively, and the intervals between periods were as shown in Table 1. No *S* was kept in the apparatus longer than 2 hr. on any day. When the interperiod interval was 0, *S*s were removed from the apparatus, taken to the adjoining room, placed in a holding cage, taken out, and returned to the apparatus in a single sequence lasting about 20 sec. While this treatment could be expected to induce some regression or recovery, it was included to keep handling effects as closely alike as possible for all groups at the close of one period and start of the next, in anticipation of the argument that handling-induced changes were not "spontaneous." When the indicated interval was 24 hr. (Ellson [1938] showed a negligible effect on a bar-press response of prolonging rest beyond this time), *S*s were removed from the apparatus, taken to the adjoining room, placed in the holding cage, and allowed to drink freely for 3 min. before being returned to their home cages. When *S* had made 10 reinforced responses in Period I, indicating that the

TABLE 1

SUMMARY OF TREATMENTS FOR THE FOUR GROUPS

Group	Period						
	I	Rest	II	Rest	III	Rest	IV
1	30 trials Left	0	40 trials Right	0	40 trials Left	24 hr.	50 trials Right
2	Left	0	Right	24 hr.	Left	24 hr.	Right
3	Left	24 hr.	Right	0	Left	24 hr.	Right
4	Left	24 hr.	Right	24 hr.	Left	24 hr.	Right

acquisition of the L response could be expected, it was assigned by reference to a table of random numbers to one of the four treatment groups, and further treatment followed accordingly.

Deprivation.—Prior to pretraining, *Ss* had been adapted to a low subsistence rate of water consumption in their home cages. After adaptation was completed and *Ss* had begun pretraining, water was at no time available in the home cage. All water was obtained either in the apparatus or in the holding cage. Just prior to Periods I and IV, all *Ss* were on 24-hr. deprivation. Since *Ss* in delayed groups were under 24-hr. deprivation at the start of Periods II or III, or both, the amount of water they had consumed prior to entry into the box had to be brought up to equality with that consumed in the previous period or periods by nondelay *Ss*. Otherwise, deprivation and delay effects would have been confounded. Therefore, .5 hr. before Periods II and III, *Ss* in delayed groups were placed in holding cages and administered drive-equating water with a graduated pipette; prior to Period II, Groups 3 and 4 were given the equivalent of 30 dippersful, and prior to Period III, Groups 2 and 4 were given the equivalent of 70 dippersful. Pilot work had indicated the possibility of response-depressing effects of temporary satiation following consumption of the entire quantity in a few seconds just prior to the period, while *Ss* in nondelay groups had consumed the same amount in a period of strenuous bar pressing; thus all but .05 ml. (equivalent of 4 dippersful) was given .5 hr. before the period, and the balance immediately before the period.

Results

Results appear in Fig. 2, 3, and 4, which present graphically the essentials of over 100 statistical tests. At the tops of the figures are indicated the outcomes of F tests of the equality of the four group means where such tests showed significance at or beyond the .05 level. Significant F's were supplemented by appropriate t's throughout, and a difference mentioned in the text should be interpreted as significant unless otherwise specified.

Relative frequency measures.—A "correct" response was defined as a reinforced response following directly the preceding reinforced response (without an intervening unreinforced response). If the period's first response was reinforced, this was also called "correct."

For a sequence of n reinforcements in which S had made c correct responses, the relative frequency of correct responses was defined as c/n, whereas the relative frequency of incorrect responses was $1\text{-}c/n$. It will be noted that this definition effectively lumps together all unreinforced responses intervening between two reinforcements. Large amounts of pilot data had shown this particular measure to be insensitive to treatment differences. Thus for periods in which the L response was reinforced (I, III) the relative frequency of L, or conditioned, responses was defined as equal to the relative frequency of correct responses. For periods in which the R response was reinforced (II, IV) the relative frequency of the L, or conditioned, responses was defined as equal to one minus the relative frequency of the correct response. This protocol was adopted to insure definition of a choice-point (the point of leaving the dipper) at which the stimulus situation could be construed as constant from trial to trial, in order that estimates of response probability in terms of the model might refer always to the same stimulus-response relationship.

Figure 2 shows the relative frequency of the conditioned (L) response in blocks of 10 reinforcements across the entire experiment. In Period I, no statistically reliable differences between groups appeared. For all groups, relative frequency was around .8 in the last block of 10, and did not approach 1 as required

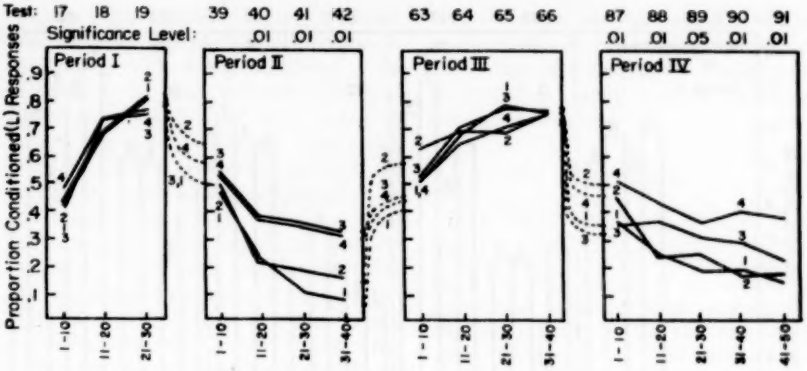

FIG. 2. Mean proportion of CR's (L) in blocks of 10 reinforcements. The dotted lines between periods represent approximate regression and recovery curves for the four groups.

for simplest application of the theory. Initial relative frequency was substantially above 0.

Between Periods I and II, relative frequency of the L response, as measured in the last 2 min. of Period I and the first 2 min. of Period II, underwent spontaneous regression as shown by the dotted lines between Periods I and II of Fig. 2. Statistical differences between groups in the amount of regression were lacking. Within Period II, a nonsignificant mean superiority of Groups 1 and 2 taken together, over Groups 3 and 4, underwent a decided reversal by the end of the period.

Over the rest interval between Periods II and III, the recovery shown by the respective groups was .1 to .4 for Group 1; .2 to .6 for Group 2; .3 to .5 for Groups 3 and 4. In Period III, no differences in relative frequency appeared at any time, all groups converging to about .75.

Spontaneous regression from this value was not uniform for the four groups—Groups 1 and 3 dropped to slightly more than .3 by the start of Period IV, and Groups 2 and 4 to about .5. Figure 2 indicates clearly that at the outset of Period IV Groups 1 and 3 together were about .15 lower in mean relative frequency of the L response, as measured along the ordinate, than Groups 2 and 4. But by the end of the period, Groups 1, 2, and 3 together fell almost .20 below Group 4. Here, as in Period II, terminal relative frequency of the L response was far above 0 in all groups, and the functions were characterized by a lack of parallelism and by intersections.

Initial values of relative frequency correct demonstrated a general increase from just less than .2 to about .6 across the entire experiment, while the terminal values stayed, on the average, between .7 and .8, regardless of period. Average initial relative frequency of the L response went from between .1 and .2 in Period I to between .4 and .6 in the next three periods; however, from Periods II to IV, it showed, if anything, a gradual decrease.

Supplementary response measures.—Figure 3 depicts the mean number of reinforcements delivered in the first 2 min. and last 2 min. of respond-

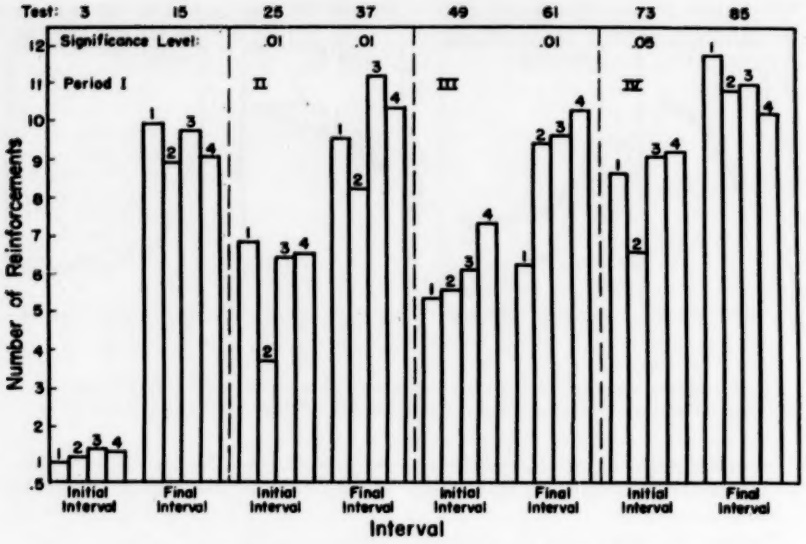

FIG. 3. Mean number of reinforcements during initial and final 2-min. intervals of responding.

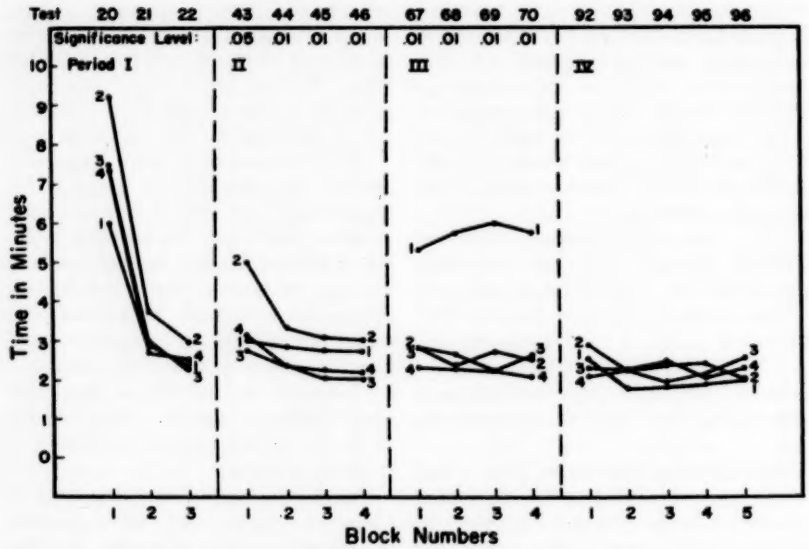
Test 20 21 22 43 44 45 46 67 68 69 70 92 93 94 95 96
Significance Level: .05 .01 .01 .01 .01 .01 .01 .01
Period I
II
III
IV
Time in Minutes
Block Numbers

FIG. 4. Mean response time by successive blocks of 10 reinforcements.

ing in each period. Figure 4 shows the time of completion of successive blocks of 10 reinforcements. In Period I, no significant differences appeared. In Period II, Group 2, when compared with the identically treated Group 1, showed significantly reduced rate of reinforcement as reflected in both Fig. 3 and 4. This occurred despite elaborate precautions to avoid such unfortunate distributions of responses, by randomizing time of day, box number, group number, etc; and a battery of tests following data collection failed to identify any accountable factor save misfortune. Had the split occurred in Period I between these two groups, the Period II measures could very probably have been adjusted by a constant amount. Adjustment of measures following Period II would, in any case, not be justifiable since identical treatments no longer prevailed. The stratagem selected was therefore simply to let the chips fall. The mean measure of Groups 1 and 2 in Period II was thus compared with that for Groups 3 and 4, notwithstanding intratreatment discrepancies. So determined, a reduced frequency of reinforcement for the immediate extinction groups (1 and 2) paralleled an elevated response time and the reduced relative frequency of the L response already noted. In Period III, response time and rate of reinforcement measures gave a clear picture—Group 4 responding at the fastest rate after having received a 24-hr. rest following each of the preceding periods, while Group 1 was markedly retarded after no rests. Rate of reinforcement was not directly related to relative frequency of the L response as it appeared to have been in Period II. Reinforcement frequency early in a period demonstrated a consistent increase across the experiment, while response times, except for the radical departure of Group 1 in Period III, exhibited a fairly regular decrease across the experiment.

Discussion

The relative frequency measure (Fig. 2) does not provide consistent support to the predictions (Fig. 1) based on the Estes model. While it is true that substantial regression and recovery, in amounts dependent for the most part upon the preceding treatment, were observed between periods, the exact nature of the dependence does not readily emerge. Perhaps the chief confirmation of the predictions lies in the fact that Group 4, which had been subjected to the 24-hr. rest after every preceding period, demonstrated a higher proportion of L responses in Period IV than either Groups 1 or 2, as predicted. And in all but the initial block of 10 trials in that period, the latter two groups were close together, as had been predicted.

However, a number of discrepancies between the predictions and results was observed, particularly in the extinction periods (II and IV). Although the initial nonsignificant superiority of Groups 1 and 2 taken together (this was depressed by deliberately introduced handling, it will be recalled) appears to confirm prediction in Period II, the reversal of this relationship by the end of the first 10 trials was unanticipated. The predicted order of relative frequencies for Period III is not in evidence, and, it should be noted, none of the obtained differences was significant. In Period IV, Group 3 was initially the lowest of the four groups, opposite to prediction, though it appeared to pick up slightly as the period progressed. In both Periods II and IV, moreover, common features are the intersection and divergence of the curves. These appear to contradict the assumption that response probabilities for all groups regardless of the length of previous

rest periods or initial differences in magnitude, should maintain a constant ordering throughout any period.

A somewhat different approach to the results might be to choose rate of reinforcement as the best indicant of response probability; that is, if reinforcement frequency per unit time were high, response probability would be interpreted as high, and so on. A formal definition of response probability on this basis would, of course, require probability to be time-dependent, as in the Poisson distribution. During extinction periods, a relatively low rate of reinforcement could be interpreted, using this approach, as indicating a relatively high resistance to extinction of the response of the previous period. So regarded, the curves for frequency of reinforcement and response time in Period II (Fig. 3 and 4) would appear to confirm predictions: Groups 1 and 2 taken together respond less rapidly than do Groups 3 and 4, presumably indicating a higher probability of the L response of Period I. This interpretation at least conforms to expectations based on conventional forgetting curves for groups tested immediately and after delay (Ebbinghaus, 1913). In Period III, the same measures indicated superiority for Group 4 and approximate equality for Groups 2 and 3, as predicted. Group 1, however, shows a response strength falling progressively below that of the groups (1 and 2) to which it should have been comparable. It could be argued here that Group 1's demise is attributable to fatigue—though initially equated with other groups for thirst drive, it remained in an apparatus where typical temperature was close to 100° F. for a longer continuous period than the other groups. Examination of Fig. 3 divulges that Group 1's response rate is not statistically different initially from the other groups, while it is clearly slower by the end of the period. Figure 4 discloses, in fact, that this response degradation is already well under way by the end of the first block of 10 reinforcements. However, fatigue is possibly confounded with additional influences here. Yet, even if a measure of success were granted to the rate-of-response interpretation of the first three periods' data, it is difficult to imagine how the outcome of this measure, as shown by Fig. 3 and 4 in Period IV, could be reconciled with the ordering of the curves in Fig. 1, for that period. Nor does it appear to the writer that any systematic application of an alternative theory, such as generalization or reactive inhibition, would help quantify the data.

Retention of relative frequency as the primary measure, on the other hand, forces attention to the disparity between the assumptions of the model and the outcome of the experiment in the particular matters of intersections of initially disparate curves, and divergence of initially equal curves of extinction. It was assumed that no intraperiod transfer of elements occurs between S and S'. Because intraperiod curves diverge so markedly in Period II (Fig. 2), the advisability of retaining this assumption is open to question. Indeed, it is this failure of curves to preserve initial relationships within a given period which does not coincide with the Estes model, and not their failure to start at $p(0) = 0$ or end at $p(n) = 1$ for the response under reinforcement. (The latter difficulty is easily overcome by substituting the empirical values of $p(0)$ and $p(n)$ for 0 and 1, an adjustment which alone cannot affect the ordering of curves.) Retention of the intraperiod no-exchange assumption precludes simple forgetting within a learning period, as well as spontaneous recovery during an extinction period, and in general imposes what appears to the writer to be a prohibitive limitation on practical application of the model.

Theoretical endeavors in psychology may sometimes have to face the impossibility of abstracting a learning situation sufficiently to permit categorical demonstrations of limited theoretical propositions. The model for spontaneous recovery and regression tested in this study apparently illustrates the point. Though the model enjoys internal consistency and a clean-cut

mathematical relation to models for other aspects of learning, it can probably be tested only in a highly restricted experimental situation—one very possibly limited to studies with man and higher primates. Furthermore, it seems only common sense to treat the no-exchange assumption as a very special case of a more general model in which exchange of elements between S and S′ proceeds both within and between experimental periods, *not necessarily at the same rate.* The development of such a model is undertaken elsewhere (McConnell, in press), and a first-order approximation to the present data is presented therein.

Summary

Data are provided concerning the influence of distributional conditions during learning on the subsequent retention of CR's. The study was designed to permit tests of Estes' mathematical model for spontaneous recovery and regression.

Ninety-six female albino rats were run in two-bar Skinner boxes. Each of four groups was subjected to four experimental periods in the sequence acquisition, extinction, acquisition, extinction. During any period, responses to one bar were reinforced and responses to the other went unreinforced. The time intervals between periods were varied among the four groups, the intervals being 0, 0, and 24 hr. for Group 1; 0, 24, and 24 for Group 2; 24, 0, and 24 for Group 3; and 24, 24, and 24 for Group 4.

It was predicted from the model that relative frequency of CR's in Period II should be higher for Groups 1 and 2 than for Groups 3 and 4; that in Period III, Group 4 should be the highest in this measure, while the other three groups should be equal; and that in Period IV, Group 3 should be highest, followed by Group 4 and then by Groups 1 and 2 together.

The relative frequency measure did not consistently demonstrate the predicted order. In Period II, this order appeared initially but differences were not significant and the order was reversed as the period progressed. In Period III, no significant differences in this measure occurred. In Period IV Groups 1, 2, and 4 appeared in the predicted order, but Group 3 was unexpectedly low. A supplementary measure, rate of reinforcement, did show the predicted order in all periods, with exceptions of a marked lowering for Group 1 in Period III, and for Group 3 in Period IV.

The discrepancies between observed and predicted orders of relative frequencies, together with systematic intraperiod changes in observed orders, question the model's assumption that spontaneous regression and recovery of the CR could be accounted for solely in terms of interperiod fluctuations in stimulus conditions. An augmented model may provide for negligible intraperiod fluctuation only as a special case.

References

Ebbinghaus, H. *Memory.* (H. A. Ruger and C. E. Bussenius, trans.) New York: Teachers Coll., Columbia Univer., 1913.

Ellson, D. G. Quantitative studies of the interaction of simple habits: I. Recovery from specific and generalized effects of extinction. *J. exp. Psychol.*, 1938, **23,** 339–358.

Estes, W. K. Statistical theory of spontaneous recovery and regression. *Psychol. Rev.*, 1955, **62,** 145–154.

McConnell, D. G. An augmented model for spontaneous regression and recovery. *Psychometrika*, in press.

(Received March 31, 1958)

Journal of Experimental Psychology
Vol. 57, No. 2, 1959

THE ROLE OF REINFORCEMENT AND NONREINFORCEMENT IN AN "APPARENT FRUSTRATION EFFECT" [1]

ALLAN R. WAGNER

State University of Iowa

Amsel (1958a) has recently presented a theory of nonreinforcement in appetitional learning situations which ascribes, under certain conditions, active motivational properties to nonreinforcement. In part, the theory maintains that after a response has been reinforced a sufficient number of times to permit the development of anticipatory goal responses, nonreinforcement will elicit a primary motivational condition termed *frustration*.

Support for this position comes largely from a series of studies (Amsel & Hancock, 1957; Amsel & Roussel, 1952; Roussel, 1952) from the Tulane laboratory. The response chain investigated involved rats running from a start box to a goal box where reinforcement was given on a portion of the trials and then running to a second goal box where reinforcement was always given. In each of these studies it was found that *after* some number of reinforcements in the first goal box, *Ss* ran faster in the second alley on trials when reinforcement was omitted from the first goal box than on trials when reinforcement was given. This increased vigor of responses following nonreinforcement has been termed the *frustration effect* (FE) and has been attributed, by Amsel, to the motivational properties of frustration.

Although the results of these studies are consistant with a theory of frustrative nonreward, Seward, Pereboom, Butler, and Jones (1957), have been unwilling to reject an alternative interpretation according to which the difference in Runway 2 running speed following reinforcement and nonreinforcement may be due, not to an increase in drive on nonreinforced trials, but rather to a decrease in drive on reinforced trials.

Seward et al. (1957) in a similar double runway situation again obtained an apparent FE. In addition, however, it was observed that *Ss* ran significantly slower in both Runway 1 and Runway 2 if they were prefed either .5 or 1.0 gm. of food before entering the first start box. Arguing that feeding in Goal Box 1 would serve to depress performance as did prefeeding, the authors concluded that at least part, if not all, of the apparent FE could be attributed to a reduction of the incentive motivational variable on reward trials rather than frustration on nonreward trials.

Some caution should perhaps be observed before one concludes that the response depressing effects of feeding will account for the apparent FE in the Tulane studies, as well as in the Seward et al. study. Amsel

[1] Part of a thesis submitted to the Graduate College of the State University of Iowa in partial fulfillment of the requirements for the M.A. degree. The author is indebted to Kenneth W. Spence for his advice and assistance throughout the course of the investigation, and to Lila M. Wagner who assisted in running *Ss*. The present report was prepared while the author was a National Science Foundation Predoctoral Fellow.

(1958b), and Amsel and Roussel (1952), have explicitly rejected such an interpretation of the Tulane results on the basis of previous prefeeding studies. These investigations (Bruce, 1937, 1943; Fredenburg, 1956; Maltzman, 1952; Morgan & Fields, 1938) appear to be in essential agreement that *small* amounts of prefeeding, if they have any effect, will lead to an increase in running speed, whereas larger amounts of prefeeding may lead to a decrement in subsequent performance. Although it is not clear at what magnitude retarding effects will appear, it should be noted that other investigators (Maltzman, 1952; Morgan & Fields, 1938) have failed to find statistically reliable depressing effects of prefeeding unless much larger magnitudes were used than those employed by Seward et al. However, granted that .5 or 1.0 gm. feeding does lead to response depression, it is doubtful that the smaller amounts of reward (.125 gm. food or two drops water) used in the Tulane experiments would lead to similar results. Furthermore, as Seward et al. have pointed out, the finding that feeding leads to response depression does not vitiate the interpretation that nonreward leads to an increment in motivation level.

A method of evaluating the frustration and response-depression interpretations would appear to be afforded by a repetition of the double runway experiment, but with an added control group which never receives reinforcement in the first goal box. According to the frustration hypothesis, since these control *S*s are never rewarded in Goal Box 1, nonreinforcement would not be frustrative. In this case the control *S*s would be expected to run slower in Runway 2 than would the experimental *S*s following a frustrative nonreward. According to the response-depression hypothesis, these control *S*s would not have their motivation level reduced by feeding in Goal Box 1 and, thus, should run faster in Runway 2 than would experimental *S*s after a reinforcement. Furthermore, if response depression is the only variable accounting for the apparent FE it would be expected that the control *S*s should run at about the same speed as would experimental *S*s after nonreinforcement in Goal Box 1.

The present investigation was designed to permit such an evaluation of the two interpretations as well as to attempt to replicate the results of the Tulane studies with regard to the development of the apparent FE under partial reinforcement conditions.

Method

Subjects

Forty-five experimentally naive female hooded rats from the colony maintained by the State University of Iowa Psychology Department served as *S*s. Their ages ranged from 90 to 100 days at the beginning of the experiment. Fifteen *S*s were randomly assigned to each of three experimental groups. Of these, two *S*s died during the course of the experiment leaving 43 *S*s from which data were collected.

Apparatus

The floor plan and dimensions of the apparatus are shown in Fig. 1. The apparatus, which is similar to that employed by Amsel and Roussel (1952), consists essentially of two runways in series, the goal box of the first runway serving as the start box for the second. The SB, Runway 1, and G_1 were painted flat black and had a rubber-mat floor. In order to differentiate the first and second alley conditions as much as possible, Runway 2 and G_2 were painted white with a hardware-cloth mesh covering on the floor. Similarly, while the goal piece in G_1 was a small plastic cup mounted on the side of the goal box, the goal piece in G_2 was an aluminum block located at the end of the goal box.

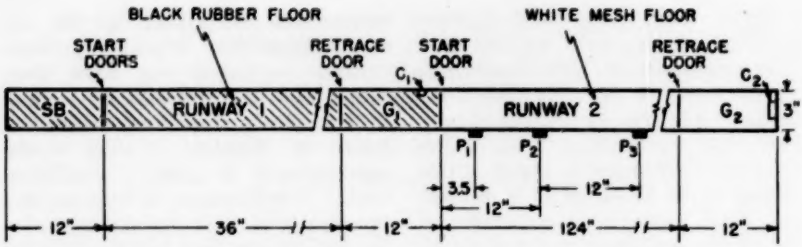

FIG. 1. Floor plan of the apparatus. The walls of the alley were 6 in. high. P_1, P_2, and P_3 represent the three photocells which were located 2 in. above the runway floor; C_1 and C_2, the two goal pieces; SB, the start box; G_1 and G_2 the two goal boxes.

Two start doors, one opaque and the other clear plastic, separated the start box from Runway 1. The single start door separating G_1 from Runway 2 was opaque, the two sides corresponding in brightness to the alleys they faced. The top of the maze was covered with clear plate glass and the goal sections with hinged pieces of clear plastic. Dim illumination was provided by the reflected light from three shielded 60-w. bulbs equally spaced 5.5 ft. above the maze.

Starting and running times in Runway 2 were obtained in the following manner. Raising the G_1 start door activated a Standard Electric 1/100-sec. clock which was stopped when *S* interrupted an infrared beam 3.5 in. from the G_1 door. A second clock was activated and stopped by the successive interruptions of infrared beams located 1 and 2 ft. from the G_1 door.

Procedure

Habituation.—The preliminary training consisted of a total of eight days of habituation during which a 24-hr. food deprivation schedule was established. The eight days consisted of: three days in which each *S* was handled for a 5-min. period before being fed for 1 hr. in an individual feeding cage; two days in which each *S* was placed in a carrying cage for 15 min. before being fed; and finally three days in which *S*s were allowed to explore the maze in groups of three for 5-min. periods. During these periods the photoelectric system was operating so that *S*s could become adapted to the clicking noise of the clocks and any other noises associated with the apparatus. After the exploration period on each of these days, *S*s were placed in an unpainted pine goal box and allowed to eat a number of the food pellets later used as reward. The *S*s were allowed to eat these pellets from the food cup, or cups, in which they would receive reward in the experimental apparatus.

Experimental procedure.—A training trial was initiated with the introduction of *S* into the starting box. As soon as *S* faced in the direction of the runway the opaque door was raised. Approximately 2 sec. later the plastic door was raised, allowing *S* to transverse the first runway to G_1. Lowering the first retrace door enclosed *S* in G_1, where it was either rewarded with a .1 gm. Purina Lab Chow pellet or not rewarded, according to the experimental plan. Fifteen seconds after the closing of the retrace door, the start door was raised, allowing *S* to enter Runway 2. The only departure from this 15-sec. interval occurred during the first two trials for the reinforced groups when it was necessary for *E* to wait until *S* had discovered the food before starting the 15-sec. period. Each *S* received two .1 gm. pellets in G_2 on every trial and was removed as soon as it had finished eating.

A total of 108 trials was given in 29 consecutive days. Each *S* received one training trial on Days 1 and 2, and two trials on Day 3. Subsequent to this, four trials were administered each day with a minimum intertrial interval of 20 min. and at an average of 22 hr. of food deprivation.

The *S*s in Group P were given partial reinforcement experience in G_1 from the beginning of training. These *S*s were rewarded in G_1 on the first three trials and nonreinforced for the first time on Trial 4. On the following trials, nonreward (F) and reward (R) were introduced in G_1 by repeating the following cycle of trials: RFRF FRFR; FRRF RFFR; FFRR RRFF. In each successive block of eight trials there were four R and four F trials, one R and one F trial appearing in each of the four daily

trial positions. The *S*s in Group P received 13 such blocks of eight trials after Trials 1–4. The *S*s in Group C were given continuous reinforcement in G_1 until Trial 77, at which time the same schedule of reinforcement and nonreinforcement as that received by Group P was initiated and continued for four blocks of eight trials. The *S*s in Group N were never rewarded in G_1.

Results and Discussion

In analyzing the Runway 2 performance, median reciprocal starting and running times were computed for each block of trials for each *S*. For *S*s in Group N this involved the computation of medians over Trials 1–4, and over each subsequent block of eight trials. The same procedure was followed for the Groups P and C measures with the exception of those blocks in which an *S* received both nonreinforcement (F) and reinforcement (R) in Goal Box 1. In these cases separate medians were computed for the F and R trials in each block. Since the starting and running measures present almost identical pictures, only the results for the running measure will be presented here in any detail.

Figure 2 presents Runway 2 mean median running speed under each experimental condition as a function of trial blocks. Considering only

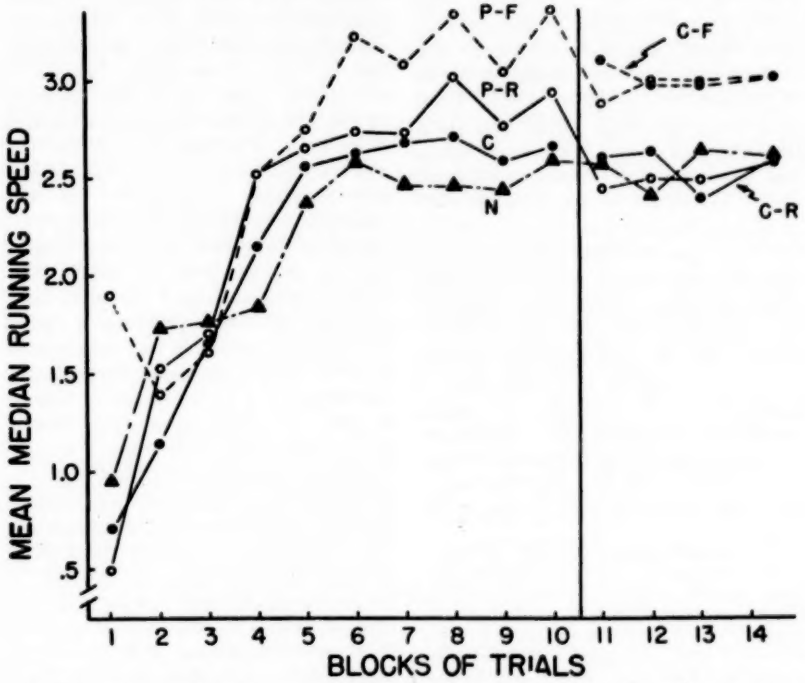

Fig. 2. Mean median Runway 2 running speeds plotted over blocks of trials for each condition. Trial Block 1 represents the appropriate trials for each condition among the first four total trials, whereas each following block represents the appropriate trials in each successive block of eight total trials. The vertical line indicates the point at which nonreinforcement conditions were first administered to Group C.

the curves for Groups P and C, it should be noted that the present results agree with those of previous studies in again demonstrating a greater asymptotic running speed on F than on R trials. Inspection of Fig. 2 also suggests that for Group P this apparent FE does not occur during the early acquisition trials but develops during training, reaching approximately the same magnitude as that which occurs in Group C immediately upon the introduction of nonreinforced trials.[2]

In order to evaluate the reliability of the apparent FE in Groups P and C over Test Blocks 11–14, in conjunction with the hypothesis that there is no difference in the asymptotic magnitude of FE between the two groups, a mixed design analysis of variance (Lindquist, 1953, Type VI) was computed. Only the data for Groups P and C were included in this analysis; Trial Blocks and Goal Box 1 Condition (F vs. R) were used as within-*Ss* factors and Groups as the between factor. The results of this analysis indicated that the only effect to attain significance was that for Goal Box 1 Condition ($F = 46.36$, $df = 1, 27$, $P < .001$). The failure of the interaction of Goal Box 1 Condition with Groups to prove significant confirms the suggestion of investigations by Amsel and Hancock (1957) with the separate reinforcement schedules, that there is no difference in the magnitude of the apparent FE following extended partial and continuous reinforcement training.

Of primary concern is the test of the response depression vs. the frustration interpretation of the apparent FE. The response-depression hypothesis interprets the greater vigor of performance following nonreinforcement compared to performance following reinforcement as being due to the demotivating effects of feeding on reinforced trials and an absence of such demotivating effects on nonreinforced trials. By the same consideration, there should be no demotivating effects upon the Runway 2 performance of Group N in the present investigation since these *Ss* are never fed in G_1. Accordingly Group N would be expected to be superior to the R-trial performance of Groups P and C. Furthermore, if the demotivating effect of feeding is the only variable responsible for the apparent FE, Group N would not be expected to differ from the F-trial performance of Groups P and C.

Inspection of Fig. 2 indicates that neither of these predictions were confirmed. Rather, in substantiation of the frustration hypothesis, the Runway 2 performance of Group N did not differ over Test Blocks 11–14 from that of Groups P and C following reinforcement, but was inferior to the performance of Groups P and C following nonreinforcement. The reliability of these differences between Group N and the two conditions of Groups P and C was evaluated by individual t tests. The results of these comparisons indicated that there was no significant difference between Group N and the R-trial performance of either Groups P or C, but that Group N was significantly below the F-trial performance level

[2] The notable exception to this statement is the performance of Group P on the first block of trials, where there exists a sizable difference between F and R running speeds. It should be noted that in this block of trials only one F trial is compared with the median of the three preceding R trials. In consideration of the large intertrial variability and the observation that no difference between F and R trials occurs in the next three blocks of eight trials, this result should probably be attributed to chance factors.

of both Groups P and C ($t = 1.40$, $df = 26, .05 < P < .10$, and $t = 1.77$, $df = 27, P < .05$, respectively). While the latter differences were only of marginal reliability, the results of similar comparisons made for the starting speed measure ($t = 2.14$, $df = 26$, $P < .025$, and $t = 2.62$, $df = 27$, $P < .01$, respectively), support the conclusion that the inferior performance of Group N, as compared to the F-trial performance of Groups P and C, should not be attributed to chance.

The present results clearly support Amsel's theory of frustrative nonreward. Considering the performance of all three groups, it is evident that behavior following nonreinforcement is more vigorous than behavior following reinforcement—but only under the condition that there has occurred previously some minimum number of reinforcements. If larger magnitudes of reward had been used, as in the Seward et al. (1957) experiment, a retarding effect of feeding might also have been observed. However, with the present magnitude (as with those in the Tulane investigations) it appears that the observed FE is independent of any demotivating effects of feeding.

Summary

Two interpretations have been proposed to account for the finding that responses immediately following the nonreinforcement of a previously reinforced response are more vigorous than responses immediately following reinforcement. The frustration hypothesis proposes that this difference in response vigor, FE, is due to an increment in drive level (D) resulting from nonreinforcement, whereas the response-depression hypothesis proposes that it may be attributed to a decrement in D as a result of reinforcement. The present investigation was designed primarily as an evaluation of these two proposals.

Forty-three female hooded rats were given 108 acquisition trials in an apparatus which consisted of two alleys in succession, the goal box of the first serving as the start box of the second. Group P was reinforced in Goal Box 1 on half of the trials from the beginning of training. Group C was reinforced in Goal Box 1 on every trial until Trial 77 after which half of the trials were nonreinforced. Group N was never reinforced in Goal Box 1. All three groups were always reinforced in Goal Box 2.

The analysis of performance in Runway 2 following reinforcement and nonreinforcement in Goal Box 1 indicated that FE was not present on the initial trials for Group P but developed with partial reinforcement training. When nonreinforced trials were administered to Group C, FE occurred on the first such block of trials and persisted with no significant difference in magnitude from that which occurred in Group P. Group N performance over the last 32 trials was not significantly different from the performance of Groups P and C following reinforcement in Goal Box 1. Over these same trials Group N performed significantly below the levels of Groups P and C following nonreinforcement in Goal Box 1. The performance of Group N relative to the other groups was considered as supporting the frustration hypothesis and contradicting the predictions of the response-depression hypothesis.

References

Amsel, A. The role of frustrative nonreward in noncontinuous reward situations. *Psychol. Bull.*, 1958, **55**, 102–119. (a)

Amsel, A. Comment on "Role of prefeeding in an apparent frustration effect." *J. exp. Psychol.*, 1958, **56**, 180–181. (b)

Amsel, A., & Hancock, W. Motivational properties of frustration: III. Relation of frustration effect to antedating goal factors. *J. exp. Psychol.*, 1957, **53**, 126–131.

Amsel, A., & Roussel, J. Motivational properties of frustration: I. Effect on a running response of the addition of frustration to the motivational complex. *J. exp. Psychol.*, 1952, **43**, 363–368.

Bruce, R. H. An experimental investigation of the thirst drive in rats with especial reference to the goal gradient hypothesis. *J. gen. Psychol.*, 1937, **17**, 49–60.

Bruce, R. H. The effects of varying the place of fractional consummatory response upon the rate of acquiring a simple learning problem. *J. gen. Psychol.*, 1943, **63**, 165–175.

FREDENBURG, N. C. Response strength as a function of alley length and time of deprivation. Unpublished master's thesis, State Univer. of Iowa, 1956.

LINDQUIST, E. F. *Design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin, 1953.

MALTZMAN, I. The process need. *Psychol. Rev.*, 1952, **59**, 40–48.

MORGAN, C. T., & FIELDS, P. E. The effect of variable preliminary feeding upon the rats' speed of locomotion. *J. comp. Psychol.*, 1938, **26**, 331–348.

ROUSSEL, J. Frustration effect as a function of repeated nonreinforcements and as a function of the consistency of reinforcement prior to the introduction of nonreinforcement. Unpublished master's thesis, Tulane Univer., 1952.

SEWARD, J. P., PEREBOOM, A. C., BUTLER, B., & JONES, R. B. The role of prefeeding in an apparent frustration effect. *J. exp. Psychol.*, 1957, **54**, 445–450.

(Received for early publication October 3, 1958)

ERRATUM

In the article "Marginal and Conditional Stimulus and Response Probabilities in Verbal Conditioning" by Jean Engler (*J. exp. Psychol.*, 1958, **55**, 303–317) the upper limit of summation in Equations 4 and 5 on pages 306 and 307 should be $t' + t - 1$ rather than $t' + t$.

The uses of apparatus and laboratory procedure in psychological experiments.

Laboratory Instrumentation in Psychology

by W. W. Grings
University of Southern California

An introduction both for the student and for the professional man or woman in behavior research who seeks detailed information and evaluation on the use of equipment and techniques.

282 pages, 142 illustrations, cloth, $4.75

Order from your bookstore or from

N-P PUBLICATIONS, *publishers of National Press Books*
850 Hansen Way, Palo Alto, California

MICROFORM

All journals published by the American Psychological Association are now available on MICROFILM or MICROCARD.

Psychological Review
American Psychologist
Psychological Bulletin
Psychological Abstracts
Contemporary Psychology
Psychological Monographs
Journal of Applied Psychology
Journal of Consulting Psychology
Journal of Educational Psychology
Journal of Experimental Psychology
Journal of Abnormal and Social Psychology
Journal of Comparative and Physiological Psychology

Available only in volume units; no single issues.

For MICROFILM, order from:
UNIVERSITY MICROFILMS, INC.
313 North First Street
Ann Arbor
Michigan

For MICROCARD, order from:
J. S. CANNER & COMPANY, INC.
Microcard Division
618 Parker Street
Roxbury 20, Massachusetts